Technische Universität Braunschweig

Gabriel Pieter Greeff

Applied Metrology in Additive Manufacturing

Dissertation
Braunschweig 2018

Berichte aus dem **Institut für Elektrische Messtechnik und Grundlagen der Elektrotechnik** Band **60**

Hrsg. Prof. Dr. Meinhard Schilling

Applied Metrology in Additive Manufacturing

Von der Fakultät für Elektrotechnik, Informationstechnik, Physik
der Technischen Universität Carolo-Wilhelmina zu Braunschweig

zur Erlangung des Grades eines Doktors

der Ingenieurwissenschaften (Dr.-Ing.)

genehmigte Dissertation

von Gabriel Pieter Greeff

aus Kokstadt, Südafrika

eingereicht am: 22.06.2018

mündliche Prüfung am: 11.09.2018

1. Referent: Prof. Dr. rer. nat. Meinhard Schilling
2. Referent: Prof. Dr. Ing. Rainer Tutsch

Druckjahr: 2018

Dissertation an der Technischen Universität Braunschweig,
Fakultät für Elektrotechnik, Informationstechnik, Physik

Kurzfassung

Die Fertigung von Produkten durch additives schichtweises Aufbauen verspricht spannende Möglichkeiten, darunter die Umsetzung von Industrie 4.0, Nachhaltigkeit, lokale Fertigung und einen Messdatensatz, der jeden Gitterpunkt des produzierten Teils beschreibt. Es werden daher Messmethoden benötigt, die einen solchen Datensatz messen und das Ergebnis interpretieren können. Eine kostengünstige additive Fertigungstechnik, die bei dem weltweiten Problem der Kunststoffverschmutzung helfen kann, ist Fused Filament Fabrication. Dies ist ein Open-Source-Konzept und ideal für die Verwendung in sich entwickelnden Volkswirtschaften. Um diese Technologie über ein Rapid-Prototyping-Tool hinaus zu einer funktionalen Produktionslösung zu entwickeln, sind jedoch neue Messansätze erforderlich.

In dieser Arbeit werden kostengünstige und leicht zugängliche Lösungen zur Verbesserung des Materialzufuhrmechanismus entwickelt. Zunächst wird ein detaillierter Überblick über die Prozessschritte gegeben, was auch zu diesem sich schnell entwickelnden Bereich beiträgt. Diese Übersicht umfasst die Bewegungswerkzeuge vom Schrittmotor-Antrieb bis zur Firmware-Implementierung, um die Motorsteuerung zu realisieren. Fünf Methoden zur Optimierung des Prozesses werden entwickelt und experimentell getestet. Dazu gehören die optische Überwachung des Fördermechanismus, die den Volumenstrom messen kann, eine Methode zur Messung der Austrittsströmungsrate, ein Drucksensor zur Messung des Verflüssigerzustands, eine Druckoptimierung mittles statistische Versuchsplanung sowie eine Verknüpfung zwischen der Messdatenverarbeitung und Auswertung.

Dies wird durch die Vorstellung des Konzeptes der physikalischen Gasphasenabscheidung weitergeführt, das dem dynamischen Schablonenlithographieprozess ähnlich ist, jedoch mit einer Konfiguration realisiert wird, die auf der Fused Filament Fabrication-Elektronik basiert. Das Design und die Konstruktion des ersten Vapor Deposition Fabrication-Mikrodruckers werden vorgestellt.Diese Arbeit schlussfolgert, dass der Materialdurchsatz in der additiven Fertigung eine wichtige Prozessvariable ist. Darüber hinaus ist eine Standardisierung wichtig für die additive Fertigung. Dies beinhaltet, wie G-Codes von der Drucker-Firmware interpretiert werden. Schließlich müssen viele Prozessparameter optimiert werden und die hier gezeigte Druckoptimierungsmethode ist ein Beispiel, das Teil einer vollständigen Druckerentwicklungs-, Qualifizierungs- und Konformitätslösung sein kann.

Abstract

The fabrication of products by additively building them layer by layer promises exciting possibilities, including realisation of Industry 4.0, sustainability, local manufacturing and a measurement dataset which describes every voxel of the produced part. Measurement methodologies, which can measure such a dataset and interpret the result, are therefore needed.

A cost effective additive manufacturing technique, which can help with the worldwide plastic pollution problem, is Fused Filament Fabrication. This is an open source concept and ideal for adoption by developing economies. Measurement approaches are needed, though, to develop this technology beyond a rapid prototyping and modelling tool, to a functional production solution.

This thesis develops cost effective, accessible solutions to improve the material feed mechanism, which is one of the most critical process components. A detailed review of the process steps is given first, which also contributes to this still rapidly developing field. This review includes the motion tool chain, from the stepper motor actuation to the firmware implementation to realise motor control. The feedstock materials and the liquefier design are also presented.

Five methods for optimisation of the process are developed and experimentally tested. This includes the optical monitoring of the feed mechanism, which can measure the volumetric flow rate, a method to measure the exit flow rate, a pressure sensor to measure the liquefier state, single print optimisation with design of experiments and a link between the in-process and post process measured data.

This is taken further by presenting the Vapour Deposition Fabrication concept, which is similar to the Dynamic Stencil Lithography process, but realised with a configuration based on the Fused Filament Fabrication electronics and firmware. The design and construction of the first Vapour Deposition Fabrication micro-printer are presented.

This thesis finds that rate material addition in additive manufacturing is an important process variable, which needs to be monitored with indirect methods. Furthermore, standardisation is important for additive manufacturing, and this includes how G-codes are interpreted by the printer firmware. Finally, many process parameters must be optimised and the single print optimisation method shown here is an example method, which can form part of a complete printer development, qualification and conformance solution.

Acknowledgements

This work was made possible by *Institut für Elektrische Messtechnik und Grundlagen der Elektrotechnik* (EMG), the Braunschweig International Graduate School of Metrology (B-IGSM) and the National Metrology Instistute of South Africa (NMISA), while the scholarship was funded by the Federal Ministry for Economic Cooperation and Development of Germany, through the Physikalisch-Technische Bundesanstalt (PTB).

I would like to thank the following people:

- Prof. Dr. Meinhard Schilling
- Dr. Harald Bosse
- apl. Prof. Dr. Frank Ludwig
- Prof. Dr. Rainer Tutsch
- Dr. Michael Martens
- Judith Krakowski and Katrin Naß
- EMG colleagues: Ralf Behme, Tanja Coenen, Georg Dornig, Kerstin Franke and Gabriela Kurbach
- EMG mechanical workshop: Harald Schmidt, Hartmut Müller and Jan Pförtner
- Master students: Uan Dung and Shihang Wang
- Fellow PhD colleagues, especially Christoph Balceris

I thank my family, parents and friends and especially my wife, Eurette Greeff. Who bravely and lovingly moved with me to a new country with a foreign language, far away from family and friends. I thank you for taking care of our daughter and supporting me throughout. It was an amazing and crazy adventure. I love you.

I honour my God, Creator and Saviour. I can only confess that to You alone belongs all the glory.

Glossary

ABS	acrylonitrile butadience styrene
AM	Additive Manufacturing
ASTM	American Society of Testing and Materials
CAD	Computer Aided Design
CMM	Coordinate Measuring Machine
DOE	Design of Experiments
FDM®	Fused Deposition Modelling
FFF	Fused Filament Fabrication
ISO	International Organization for Standardization
IRM	Information Rich Metrology
ISR	Interrupt Service Routine
JTAG	Joint Test Action Group
LCVD	laser-assisted chemical vapor deposition
ME	Metrial Extrusion
NC	Numerical Control
NIST	National Institute for Standards and Technology
NMI	National Metrology Institute
SMD	Surface Mount Device
PC	polycarbonate
PCL	polycaprolactone
PLA	poly(lactic acid), also known as polylactide
PLD	pulsed laser deposition

PVA	polyvinyl alcohol
PVD	Physical Vapour Deposition
PWM	Pulse Width Modulation
RADDS	RepRap Arduino-Due Driver Shield
Raps128	RADDS Power Stepper Driver
RepRap	Replicating Rapid-prototyper
RTD	Resistance Temperature Detector
ROI	Region of Interest
SPO	Single Print Optimisation
TPE	thermoplastic elastomers
TPU	thermoplastic polyurethane
TPC	thermoplastic copolyester
VDF	Vapour Deposition Fabrication

Symbols

A	m^2	area
a	$\mathrm{mm/s}^2$	acceleration
α	rad	step angle
α_d	m^2/s	thermal diffusivity
C_{pulley}	mm	circumference of pulley
Δv	mm/s	speed difference
E	Pa	Young's modulus
η	$\mathrm{Pa} \cdot \mathrm{s}$	viscosity
ϵ	1	emissivity
F_{max}	N	maximum force
f_{FPS}	frame/s	frame rate
f_v	step/s	step pulse rate
h	mm	height
J	$\mathrm{kg} \cdot \mathrm{m}^2$	inertia
λ	mm	mean free path
m	kg	mass
n_{spr}	fullstep/rev	fullsteps per millimetre
ω	rad/s	rotational speed
P	Pa	pressure
$\dot{Q}$	$\mathrm{mm}^3 \cdot \mathrm{s}$	volumetric flow rate
$\dot{Q}_c$	W	conduction heat flow rate
$\dot{Q}_r$	W	radiation heat flow rate
R	mm	radius
R_{rod}	K/W	heat flow resistance
S	mm	distance
S_{cnv}	μstep/mm	step conversion factor
S_μ	μstep/step	microsteps per full step
T	$^\circ$C	temperature
τ	$\mathrm{N} \cdot \mathrm{m}$	torque
θ	rad	angle
φ	$\mathrm{m}^{-2} \cdot \mathrm{s}^{-1}$	evaporation flux
v	mm/s	speed
$\upsilon_{poisson}$	1	Poisson's ratio
$\dot{\gamma}$	s^{-1}	shear rate

Contents

1 Additive Manufacturing in Context of Metrology

Additive Manufacturing (AM) is the "process of joining materials to make parts from 3D model data, usually layer upon layer, as opposed to subtractive manufacturing and formative manufacturing methodologies" [1]. AM is a rapidly growing manufacturing method [2], which is emphasised by considering three perspectives on the significance of AM, namely: growth of the (1) market, (2) number patents and (3) research output as shown in Fig. 1.1.

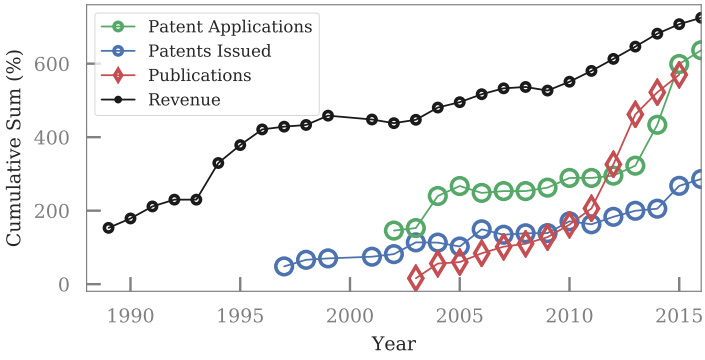


Figure 1.1: The year-on-year cumulative growth of AM, in terms of patents and revenue [2], as well as the number of publications [3].

The cumulative percentage growth of the AM revenue, number of patents (issued and applications), and the number of publications per year is shown in Fig. 1.1. This was calculated by adding the year-on-

year percentage growth for each measure. All three measures present an exciting growth curve. For example, the number of patent applications in 2001 was 24 and in 2016 it increased to 1842 [2]. The review by Baumann and Roller [3] finds that the average number of published works increased by 41.3 %, with approximately a 1000 publications in 2002 and more than 4000 in 2016. The number of desktop printers sold per year also increased from 66 in 2007 to more than 42 000 in 2016 [2]. This indicates that the investment of both industry and academia in AM is rapidly growing.

1.1 Additive Manufacturing, Industrialisation and Metrology

AM is an important technology for the realisation of *Industry 4.0*, which is said to be the fourth industrial revolution [4]. Industry 4.0 can be defined as smart and sustainable factories, which profits from the cooperation of cyber-physical systems with the Internet of Things (IoT), as well as the redefinition of the human role in manufacturing [5].The successful implementation of Industry 4.0 is essential for a strong competitive position of a country, but also provides a possible solution for global challenges, such as resource and energy efficiency [4].

A comparison between the time lines of the definition of the metre, dimensional measurement technology and milestones in industrialisation are presented in Table 1.1, which links industrialisation with metrology, the science of measurement.

Mechanisation is seen as the first industrial revolution, with the introduction of water and steam powered manufacturing. An example of this is the mechanical looms built in 1784 [4]. Simultaneously, the development of manufacturing processes rely on the existence of accurate dimensional metrology, which was made possible with definition of the decimal system in 1799 and the first micrometer, which was already produced in 1848 [6, 7].

Electrical power enabled the second industrial revolution. In 1870 the first production lines were used in Cincinnati slaughterhouses [4]. During the same period the Metre Convention was signed in 1875 and gauge blocks were developed in 1896 [7].

Gauge blocks were invented to address a significant hindrance for the global adoption of industrialisation. The production and use of interchangeable parts was made possible with water, steam and electrical power, but these parts required numerous gauges to check each part. This also necessitated a lot of effort and cost to calibrate each unique gauge. The issue was solved by C.E. Johannson in 1896, by producing a set of combinable gauges which could be joined to form various lengths, significantly reducing the number of required unique gauges. Furthermore, these gauges were combined by simply wringing them together. Consequently, gauge blocks became extremely popular and are even still in use today. In America, Henry Ford adopted the use of gauge blocks and it soon became the primary length transfer standard [8].

Table 1.1: Time line comparison between the development of the metre, dimensional measurement instruments and the three industrial revolutions (references in text).

Year	**Metre Definition**	**Dimensional Measurement**	**Industrialisation Examples**
1784			Mechanical loom
1799	Decimal Metric System		
1848		Micrometer	
1870			Production lines
1875	Metre Convention		
1896		Gauge Blocks	
1957		CMM	
1960	International System of Units (SI)		
1960	Krypton 86 radiation		
1964		He-Ne length standards	
1969			PLC, Modicon-084
1983	Speed of light		
1996		X-CT for manufacturing	
2018	SI redefinition		

Industrialisation was made possible through interchangeable parts, which could only be interchangeable if they were within specified tolerance limits. Conformance testing of the parts was therefore required, which necessitates accurate and cost effective measurements, according to international agreed physical units and standards. This is only possible with metrology, which provides an unbroken chain of measurements to the international standards.

Later, in 1957, the first universal Coordinate Measuring Machine (CMM) was introduced by the Moore Tool Company [9]. The definition of the metre was changed from a line scale and redefined based on the wavelength of Krypton-86 radiation, along with the definition of the SI, the international system of units, in 1960. The stabilised Helium-Neon laser also became an important length standard around 1964 and the definition of the metre was subsequently changed to be based on the speed of light, an universal constant, which pre-empted the expected SI redefinition in 2018 [7].

Electronics and IT (Information Technology) created the third industrial revolution, which started around 1970. An example of an enabling technology for the third industrial revolution is the Modicon 084, which was the first programmable logic controller (PLC) [4].

1.2 Additive Manufacturing and Industry 4.0

The successful integration of AM into commercial manufacturing is an essential part of Industry 4.0 [5], which raises the question of how AM will be integrated into industry.

Three scenarios are sketched in [10] for the diffusion of AM technology into the market (see Fig. 1.2): (1) adoption by the commercial manufacturing industry, (2) the *copy shop* model and (3) end-user production.

End-user production is the most extreme case, in terms of manufacturing change. This is when the consumer prints a significant amount of parts at home. The second scenario, the copy-shop model, is where part files are sent to a local business, which manufactures it. Both of these are seen as less likely. A possible fourth technology diffusion model is also mentioned, which is the use of AM for centralised spare part production.

This can be discussed further by considering the supply chain. Each partner in the traditional manufacturing supply chain must keep an inventory of both raw and process parts. A lean supply chain aims to reduce the cost and risk of such a stockpile, as much as possible. AM, however, can make this cost obsolete [10]. A conceptual comparison of the supply chain difference between AM and traditional manufacturing is shown in Fig. 1.2, where each arrow indicates a logistics solution, i.e transportation of goods, and the small grey containers the need for storage or inventory.

The material suppliers in the AM supply chain are also closer to the consumer, exposing them to a more complex relationship, than with the traditional supply chain [11]. This complex relationship is depicted with the intersecting arrows in Fig. 1.2. The material suppliers must also

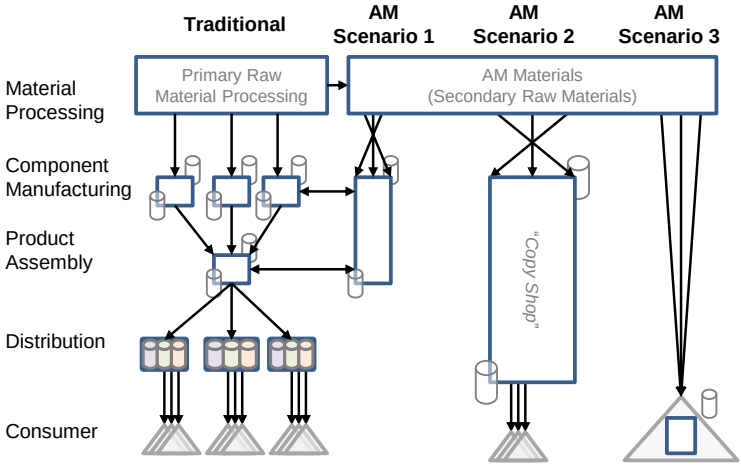


Figure 1.2: Comparison between the distribution chain of traditional manufacturing and AM.

develop processes to refine materials for a specific AM process and part requirements.

A review by Tofail *et al.* [12] finds that AM feedstock materials have more design requirements than conventional manufacturing materials, since AM feedstock is seen as a secondary raw material, requiring careful pre-processing, before AM processing.

The commercial success of AM, according to Tofail *et al.* [12], is dependant on the four interdependent factors called the four *M*s: market, making, materials, and metrology. AM processes (make) must be cost-competitive (market), producing functional and useful parts (market and make), which are according to a standard, reproducible and within tolerances (metrology), which is only possible with specially designed feedstock (materials).

Raw materials must be processed into standardised feedstock, concurrently designed with metrology, the specific AM process and market needs in mind. Apart from the required metrology for the material qualification, traceable in-line metrology is also essential for Industry 4.0 [13] and AM [14]. Many general AM measurement and design principles can be applied to all AM processes. Each AM process, however, also has requirements specific to itself.

1.3 Additive Manufacturing Processes

Seven AM process categories are identified in the international standard [1] and given in Table 1.2. Each category is realised with different tech-

niques and most of these technologies are still rapidly developing. A detail comparison is therefore difficult. There are also new processes in development. Nevertheless, the minimum layer height and possible processing materials are included in the table to give a general idea of the possibilities within a category. The following paragraphs shortly introduce each process and discuss general advantages and disadvantages, based on the reviews by [12, 15, 16].

Table 1.2: The seven AM categories, with possible feedstock materials and minimum layer height [1, 16, 17].

Process Category	**Materials**	**Layer Height** (µm)
Binder Jetting	Metals, Polymers, Ceramics	90
Direct Energy Deposition	Metal powder and wire	30
Material Extrusion	Polymers, Metals	100
Material Jetting	Photopolymers, Wax	13
Powder Bed Fusion	Metals, Polymers, Ceramics	30
Sheet Lamination	Hybrids, Metals, Ceramics	50
Vat Photopolymerisation	Photopolymers, Ceramic	25

Binder jetting (BJ) is the AM process in "which a liquid bonding agent is selectively deposited to join powder materials" [1], where the bonding agent is printed with inkjet print heads. Synonyms or trade names for this category include 3D-printing (3DP) and ink-jetting. It is a relatively fast process, with a wide material selection, which provides design freedom. The produced parts are, however, fragile.

In Directed Energy Deposition (DED) "focused thermal energy is used to fuse materials by melting as they are being deposited" [1]. An advantage of this process is the mechanical material properties of the produced parts. It is also well suited for the repair of worn-out parts or for adding a special material layer to an existing shape for enhanced product functionality. Trade names or technology terms for this category include Laser Engineered Net Shaping (LENS) and Electron Beam Welding (EBW).

Material Extrusion is defined as the "process in which material is selectively dispensed through a nozzle or orifice" [1]. This technology category can be realised with inexpensive machines and is well suited for multi-material printing. Part geometry is, however, limited by poor

surface finish and is also susceptible to vertical anisotropy. This process is reviewed in Chapter 2.

Material Jetting (MJ) selectively deposits droplets of material to build a three dimensional object. Example materials include composites, ceramics and biological material. A disadvantage of this category is low part strength, while advantages include the high accuracy of the droplet deposition. MJ can be realised with Polyjet, 3D Inkjet or Direct Ink writing.

The Powder Bed Fusion (PBF) is an AM "process in which thermal energy selectively fuses regions of a powder bed" [1]. PBF technologies include Selective Laser Sintering (SLS), Selective Laser Melting (SLM), Direct Metal Laser Sintering (DMLS) and Electron Beam Melting (EBM). In this process the powder bed acts as an integrated support structure, providing significant design freedom. Post-processing powder removal can be a challenge. Nevertheless, fully dense parts are achieved with PBF, with high specific strength and stiffness.

Sheet Lamination (SL) is the "process in which sheets of material are bonded to form a part" [1]. Metal objects can be produced in this category with Ultrasound Consolidation (UC) or with Ultrasound Additive Manufacturing (UAM). Laminated Object Manufacturing (LOM) are used for rapid prototyping, where paper like sheets are cut with a laser beam and laminated together. Advantages include high speed and low material cost. The strength of the final LOM part, however, depends on the adhesive used and this also limits the number of different materials which can be used.

Vat Photopolymerisation (VP) is defined as the "process in which liquid photopolymer in a vat is selectively cured by light-activated polymerization" [1] and is also the first AM process. The development of this process can be traced back to the 1960s, when the first attempt was made at the Battelle Memorial Institute to create solid objects with photopolymers using lasers. Two intersecting laser beams with different wavelengths were used to attempt the solidification of the resin through photopolymerisation. Stereolithography (SLA), the single laser beam approach, was only developed in the 1980s. The experiments, published in 1981 by Kodama [18] of the Nagoya Municipal Industrial Research Institute, are considered as evidence of the first working AM machine [19].

Example technologies for VP is SLA and Digital Light Processing (DLA). Advantages include the possibility of producing large parts, higher resolution, as well as excellent surface finish. This category is however limited to photopolymers, which have poor mechanical properties and a relatively short shelf life.

1.4 The Additive Manufacturing Process Chain

The general AM process chain is presented in Fig. 1.3, based on [20, 21]. This figure not only shows the standard steps of Computer Aided Design (CAD), file conversion, slicing, printing and post-processing, as discussed by Gibson, Rosen, and Stucker [20], but also includes additional

steps, which form part of the whole process. Each process category discussed previously requires a different solution for each step, but the process chain can still be generalised for AM. Each step in the process adds cost, risk and variance. This can be modelled with a *Ishikawa* diagram to find the final part influence factors [22].

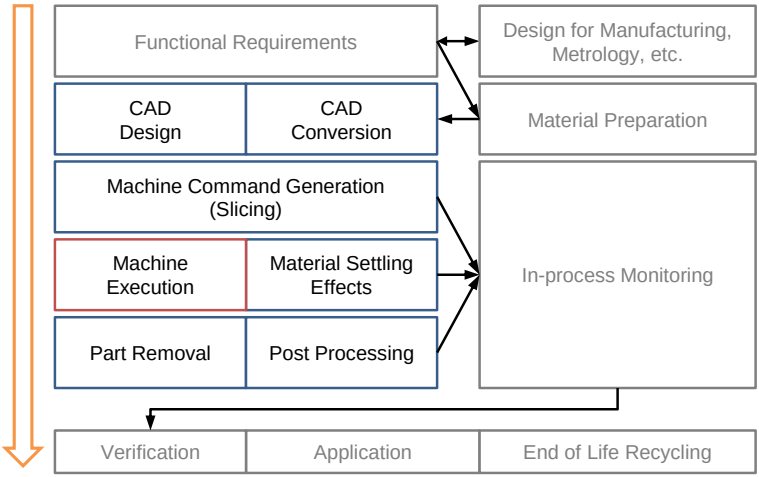


Figure 1.3: The general AM process chain. Lighter (grey) text indicate standard manufacturing steps and darker (black) AM specific steps.

The CAD and 3D modelling of objects are key elements of the AM process; as matter of fact, AM will be impossible without 3D graphics. The CAD file can either be made from scratch with a CAD program or can be created based on data from a 3D-scanner. The file must be converted into a standardised and slicer friendly file format, for example STL (originally stereolithography, but recently Standard Triangle Language or Standard Tessellation Language) [23]. The slicer is responsible for combining the machine process model with the supplied CAD file and producing a command file for printing of the object.

Functional requirement definition, a standard engineering step, is also included in the graph. Such requirements must be translated to measurable feature specifications and tolerances. This will influence both the manufacturing process and measurement tool selection.

AM provides new opportunities and challenges for metrology in this context, since part specifications can potentially be measured during the manufacturing processes. This provides an unparalleled amount of part data and the challenge is to link this to the part specifications and functional requirements. In principle, this should reduce post processing, quality assurance checks, as well as provide access to features, which are difficult to reach. Detailed information per product is measurable with

AM and this fits well with the Industry 4.0 ideals, but also asks metrology for methods to provide confidence in these measurements.

Metrology in AM is not only concerned with the geometric dimensions and tolerances, but also the material properties. This includes, for example, powder particle size measurement and purity. In-process parameters like time of exposure to a thermal source creates different part mechanical properties, which also needs to be modelled, controlled and verified [12].

Material settling effects occur during and after the execution of the print commands. The final product therefore not only depends on the exact execution of the commands by the printer, but also depends on effects such residual internal stress, due to thermal deformation.

Most AM processes rely on a phase change of the material to achieve desired object geometry. This phase change is, however, very inhomogeneous (on purpose) and localised, which complicates the modelling, as well as the interpretation, of the in-process measurement results.

The object must be removed from the build area after print completion, after which some AM processes also require post processing. Finally, the object can be used for the intended application, after the final inspection. Consideration for the end of the product life should, however, not be neglected [24].

1.5 Metrology in Additive Manufacturing

AM presents great opportunities for the aerospace industry and product development for space application pushes manufacturing technologies in general to the edge [21]. It can therefore be expected that the space application of AM will drive the developing of it, in terms of accuracy and reliability. The industry review by Dordlofva and Törlind [21], however, finds that AM still needs to build up industry confidence, with product qualification, for example. Successful part qualification proves that the design intent is fulfilled. It is found that this is still required for AM parts, as well as a link between the AM process, part geometry, material and mechanical properties [21].

The review by Tofail *et al.* [12] notes that quality assurance is not only important for improving the technology, but it is also essential to satisfy the market pull for reliable product performance. This could soon become a legal requirement. Process and product qualification require measurements with defined confidence intervals, which can only be achieved with metrology.

Note for example, the ISO definition of an AM part [1]:

> part — joined material forming a functional element that could constitute all or a section of an intended product. Note 1 to entry: The functional requirements for a part are typically determined by the intended application

As well as the definitions of an AM part layer: “layer <matter> material laid out, or spread, to create a surface” [1]. An AM part is built to fulfil a function, which it realises by joining surfaces. Surface metrology

is, therefore, a clear requirement. This can also be expanded to the functional level, which means that metrology for AM must be able to measure the functional elements.

1.5.1 Requirement for Measurement Traceability

A measurement requires an unbroken chain of measurements with estimated uncertainties, all the way back to the SI definition of a unit, for example, the metre. This is called measurement traceability and is depicted in Fig. 1.4 and links the object being measured (the work piece), through working and transfer standards to the national realisation of the metre at the local National Metrology Institute (NMI). The local realisation is based on the SI definition and prescribed guidelines for the method of realisation.

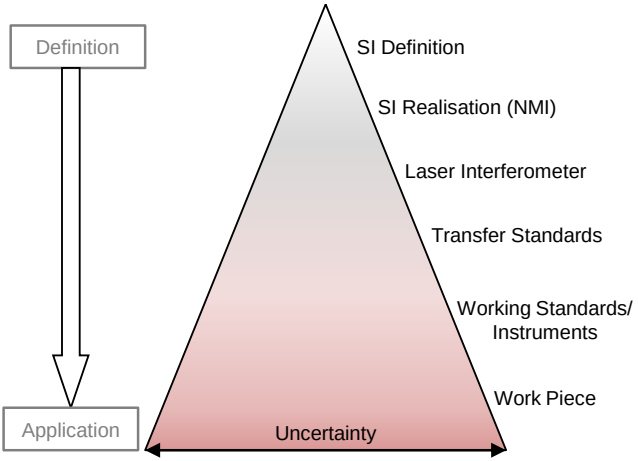


Figure 1.4: The conventional traceability chain for the metre, starting from the SI definition to the work piece (application). The measurement uncertainty increases as the traceability chain becomes longer.

The uncertainty of the measurement increases with the length of the traceability chain. The width of application, however, also increases, for example, a stabilised frequency for the national metre standard can (only) be used to calibrate the laser frequency of an interferometer. A laser interferometer can be used to calibrate a CMM, which can be used to measure almost any 3D object.

The need for metrology in AM is on the agenda of many national and international initiatives. The "*Measurement science roadmap for polymer-based additive manufacturing*" by the National Institute for Standards and Technology (NIST) [25] identifies metrological barriers for the adoption of

polymer-based AM. A need for faster and higher resolution measurements is found, as well as the need for technologies to deal with *big-data*, which will be the result of fast, high resolution sensors.

The measurement science roadmap sets a 10 µm spatial resolution target for thermal, density, chemical and dimensional measurement, as well as the requirement of a voxel based history of the entire build volume. Furthermore, the report presents the concept of a digital thread, which links process modelling, material characterisation and in-process measurement with the final performance. Input into this thread is the CAD model, material, process and equipment uncertainties, which should be used in this digital thread to predict variance of the final object.

It is possible to measure the printed part on- , in-, at- and off-line [12]. In-line metrology follows the process at the solidification point, for example at the melt-pool or deposition location. On-line metrology is used to keep track of a voxel, after creation, building a history of the point. This can be used to monitor material settling effects or consolidation. If a sample is removed during production then at-line metrology can be used. The final product is inspected for conformity to the set tolerances with off-line metrology. Fast, non-contact, non-intrusive measurement techniques are required for the in-line and on-line measurements, with for example optical and thermal camera systems. Slower methods, on the other hand, such as contact probe CMM or AFM can be used at-line or off-line [12].

A spatial resolution of 10 µm will require at least 1×10^{12} measurement points for each measurand, in a cubic build volume with 100 mm sides. Additionally, the resulting dataset needs to be interpreted, in order for it to be useful. The derived information should ideally be used in-process for real time correction and not just monitoring.

1.5.2 Information Rich Metrology

On the one hand high resolution is required, on the other hand the data must be measured quickly and remotely in order not to impede into the manufacturing process. A possible solution to this conundrum is presented by Leach *et al.* [26], who calls it Information Rich Metrology (IRM). The goal of IRM is to combine information known before the measurement to overcome the range (or speed) versus resolution limitations. AM provides a clear opportunity and need for IRM. The CAD, material, process and sensor models can be combined to overcome these challenges [26].

The measurements must also be traceable, which requires uncertainty estimations for each point. Machine tool calibration has been studied and a review of the traceability of on-machine tool instruments are presented by Mutilba *et al.* [13], who also present the idea of fast feedback of in-process or on-machine measurements, to improve manufacturing processes and reduce waste.

1.5.3 Additive Manufacturing Conformance Assessment

The concept of metrology in AM with conformity assessment, which is the "activity to determine whether specified requirements relating to a product, process, system, person or body are fulfilled" [27], is presented in Fig. 1.5. A selected feature must be designated to be within a set tolerance, which defines the permissible values for the property, called the measurand. The process used to confirm where the property is within the tolerance, by including the measurement uncertainty of the property, is called conformity assessment or inspection (note: the term *conformity assessment* here has different meaning than that used in statistics [27]).

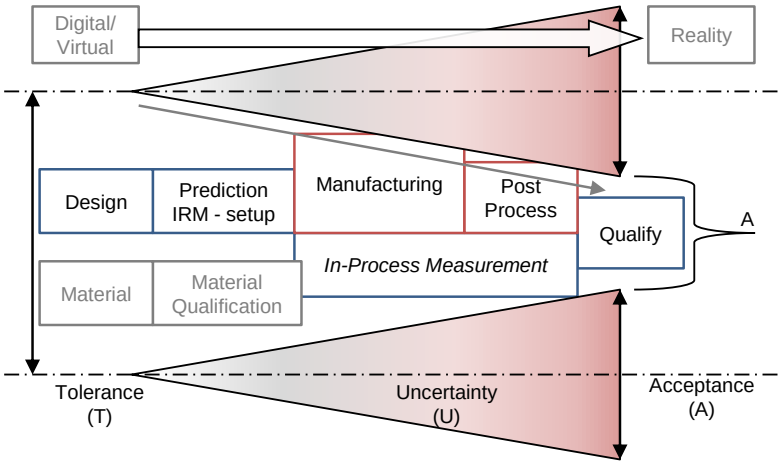


Figure 1.5: Metrology in AM with in-process measurement using Information Rich Metrology (IRM) and conformity assessment. The dash-dot lines indicates the tolerance interval (T), as set by the design. Various models and the design is used to set up the IRM model. The object is transformed from a digital representation to a solid surface through the AM process, which increases the object deviation, whilst in-process measurement and post-processing are used to ensure the measurand is within the acceptance interval (A) with measurement uncertainty (U).

The basic principle is that an acceptance interval is created by the combined measurement uncertainty probability and the tolerance interval. Conformity, as treated as binary decision (accept/reject), can only be proven if the measurand lies within the acceptance interval.

It is attempted here to combine all the above concepts, including IRM, traceability and the general AM process (Fig. 1.3), into Fig. 1.5. First, an object is concurrently designed, with the selection of materials and the fabrication process, which results in a digital representation of the object.

This object has a small or defined to be zero uncertainty, similar to the conventional traceability chain shown in Fig. 1.4.

The design is combined with the CAD model, material properties and process models, as well as the models of the in-process sensors. This creates a predicted object, with an estimated variance. The next step starts the manufacturing process. Here the in-process measurements are combined, in real time, with the models to speed-up the measurement process and increase the measurement range with IRM. The measurement output, as related to the machine input trough the process model, can be used for *on-the-fly* process parameter improvement, error correction or early discontinuation.

The in-process measurement is continued during post processing, where more time-consuming measurements can be performed. The results are, however, always considered in the framework of the feature requirement, derived tolerance and estimated measurement uncertainty. Finally, the product is checked for quality control purposes and verified to be within or outside the acceptance interval.

The in- and post-process measurement data are used to iteratively modify the various models, as well as the design rules. The next part will, however, not necessarily be the same as the previous part, requiring that the process itself is modelled accurately and any defects ascribed to root causes.

In principle, it is now possible to have a complete record of the part, from functional requirements, design, feedstock material, in-process measurement to the final part qualification. Furthermore, measurement traceability and part qualification can be estimated by using the principle of conformity assessment.

Currently, many cutting edge, factory ready, metal AM machines are only starting to introduce in-process monitoring features [14] and the metrology required to handle the data and sensing techniques are also relatively new [26]. This creates an exciting time for technological development, which combines diverse research fields.

Many questions are still open, for example, in the survey of metal AM for space applications, one correspondent compared AM to welding and noted the failure of research to overcome welding process defects [21]. This simply asks if AM will provide more reliable results than 3D welding? The answer seems to be that this is only possible if metrology can provide the required quality assurance for the market, measurement accuracy for technology improvement and fast measurement systems for real time control.

1.6 Thesis Layout

This thesis focusses on molten material extrusion with Fused Filament Fabrication, with the aim of using Fused Filament Fabrication (FFF) process understanding and improvements to build a micro resolution Vapour Deposition Fabrication (VDF) 3D printer, based on dynamic stencil lithography.

The challenge is to combine the two seemingly diverging topics, i.e. the improvement of the FFF extrusion process and the VDF realisation. This is addressed by realising that the VDF process uses the same software and electronics base as the FFF process.

This work concentrates on the *in-process measurement* block, as shown in Fig. 1.5. A *bottom-up* approach is used, by considering one of the foundational building blocks of the FFF process — the deposition of a single track, which builds layers and finally the printed part. All the steps and components required for the execution of a single FFF extrusion command are reviewed in Chapter 2. Once this foundation is laid, improvement of the FFF extrusion is presented in Chapter 3. This leads to development of the Single Print Optimisation (SPO) method, presented in Section 3.3, which can be used to quickly test and improve the VDF process. Additionally, useful contributions are made in the field of FFF manufacturing in terms of understanding the extrusion process, which can in turn be used for future process improvements. This is also significant, since FFF is such a rapidly growing, but open source, field. This means that there are not many works, which gather all the components required for extrusion in one report, this includes G-code parsing, stepper motor real time motion planning, material properties and extruder design.

The VDF chapter briefly introduces the concept and builds the system based on the FFF process components. The two biggest changes are the addition of a vacuum chamber and the material source. A purpose built resistive heater evaporation source is presented and first VDF test results.

1.6.1 Thesis Questions

The work is kept in context by the *thesis questions*. These questions are derived from the discussion in this chapter, which relates AM to metrology and industrialisation.

The questions are summarised as follows:

1. If Additive Manufacturing (AM) is an industrial revolution, does it require *new metrology*?
2. Can in-process metrology be used to improve a desktop, open source, AM Fused Filament Fabrication (FFF) process?
3. Is Vapour Deposition Fabrication (VDF) feasible, using a Fused Filament Fabrication (FFF) base?
4. Can the answers found in Question 2 and 3 be made universally applicable to AM, by providing additional insight into Question 1?

2 Review of Fused Filament Fabrication

The Material Extrusion AM process category is realised with different techniques, for e.g. FFF, robocasting, Fused Layer Modeling or Fused Layer Manufacturing. Another common name used in literature is Fused Deposition Modelling (FDM®), which is a registered trademark of Stratasys Inc.; FDM® was invented by Scott Crump, who received a patent for it in 1992 [28, 29]. This patent expired without being renewed. This created the opportunity for the Replicating Rapid-prototyper (RepRap) movement which was founded in 2004 [16, 29, 30].

A RepRap 3D printer "is an open-source self-replicating rapid prototyping machine"[30], based on FFF. The term FFF was also a created by RepRap project members to avoid confusion with FDM® [30] and to ensure freedom from trademark and patent restrictions. The RepRap open-source community is arguably responsible for the rapid growth and popularization of 3D printing. There are already 75 official RepRap variants [31] (2018). This led the foundation of many start-ups, which either designed new FFF printers or focused on the supply of components, including the production of filament.

Higher resolution parts with improved densities and mechanical properties can be achieved with other AM process categories [32]. Another disadvantage of the FFF process is part vertical anisotropy [12]. Nevertheless, advantages of FFF include inexpensive system cost, multi-material processing possibilities, strong potential for in-space applications, scalability and functional part manufacturing ability [12, 15, 16]. The open source movement behind low cost FFF systems also creates opportunities for research which are unrestricted by the intellectual property barriers of other processes. Open-source can lack standardisation and special care for developing standards across the community is required.

The FFF process is a sub-category of material extrusion AM. Paste material extrusion is another process, which is also categorised in this

group. This review will, however, focus on the FFF process. The whole AM process from concept to product is also not reviewed here, but rather the realisation of the print command. This review rather takes the *bottom-up* approach to the printing process. It considers the final part as result of many sequential extrusions or print commands, and assumes that the whole print result depends on this sub-process.

Topics such as path generation by the slicer algorithm (pre-print) or deposited track bonding and settling effects (post-print) are not reviewed in detail in this chapter. Both the pre-print and post-print processes, however, depend on how well the commanded extrusion is executed. If this is not the case, then optimisation of the pre- and post- processes are impractical. On the other hand, the review of the extrusion process is useful for understanding the pre- and post-extrusion process requirements. This review is also used extensively in the FFF extrusion improvement chapter (see Chapter 3).

This work is based on the RF1000 printer sold by Conrad Electronics [33] in 2015, which is a single extruder FFF printer with a heated bed. Components of this printer are introduced throughout the text, as needed, to explain or simulate certain concepts. These principles can, however, be generalised to most FFF printers.

An overview of the process, hardware, electronic and auxiliary components is presented first, followed by a short review of multi-material and multi-extrusion fabrication. The mechanical drive train, G-code interpreter and stepper motor motion actuation are presented next. These concepts are then used in the review of the real-time velocity planning used by FFF printers, which show how a G-code command is translated into physical motion. Feedstock materials are reviewed next, with specific focus on poly(lactic acid), also known as polylactide (PLA). This material is used to simulate the extruder models in the following section, which describes the liquefier pressure drop, nozzle design and feed slippage, amongst others. Finally, overviews of FFF process monitoring, control and optimisation are presented.

2.1 Fused Filament Fabrication

The International Organization for Standardization (ISO) defines material extrusion as an "additive manufacturing process in which material is selectively dispensed through a nozzle or orifice" [1]. Thomas and Gilbert [10] adds: "these machines push material, typically a thermoplastic filament, through a nozzle onto a platform that moves in horizontal and vertical directions". An FFF system is differentiated from other material extrusion processes, by melting the thermoplastic feedstock material in a liquefier.

The general AM process flow from part specification, through CAD generation to the final printed object, as presented in Section 1.4, is also applicable to an FFF printer.

A program to convert the CAD object to machine understandable G-code commands is required. FFF printers use a standardised program-

ming language called G-Code, which is also known as RS-274. This is widely used to control machine tools, including AM machines [34–36]. The G-codes are either read from a storage device (for e.g. a SD card) or sent in real-time by a host computer to the printer. The commands are interpreted and executed by the firmware installed on the main microcontroller. The microcontroller completes the given command by controlling the various sub-systems through the printer electronics.

The FFF process, specifically the execution of the G-code command, is reviewed next by considering the main process, hardware and electronic components, as well as auxiliary hardware components.

2.1.1 Process Components

The main components of a FFF printer are shown in Fig. 2.1, based on [28, 30, 31]. The feedstock material is pulled out of the supply spool, with the material feed mechanism. This is usually a cost effective pinch-roller type drive [30]. The feed mechanism forces the melted material through the nozzle, which has a smaller diameter than the filament. Displacement of the build bed or print head during the extrusion process allows for the deposition of object layers in the XY plane, by the selective sequential extrusion of tracks or roads.

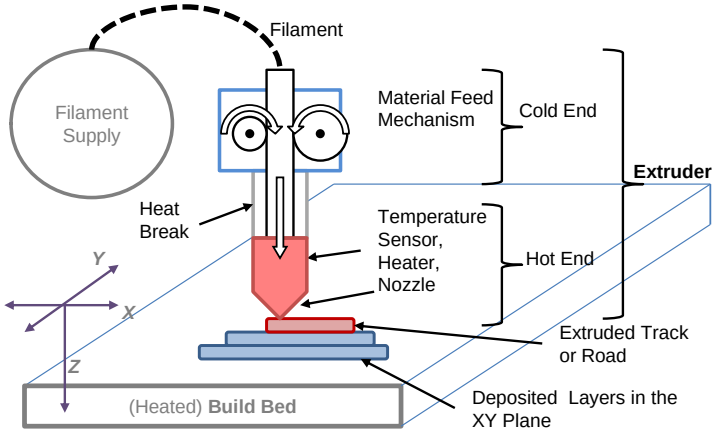


Figure 2.1: Main components of the FFF extrusion process.

The most important component of all the main components is the extruder, according to the inventors of the RepRap concept Jones *et al.* [30]. The material liquefier, combined with the feeding mechanism, is called the extruder (see Section 2.8.1 for other feed mechanisms), while feedstock material is called *filament*. This work will focus on conventional, filament based, FFF printers.

The extruder has three main components (see Fig. 2.1), namely: the cold-end (material feed mechanism), hot-end and heat break. The hot-end consists of a nozzle, heater and a temperature sensor, while the function of the heat break is to ensure that the filament temperature at the feed mechanism remains below the glass transition point. This ensures that the feed mechanism can grip the material and reduces the risk of filament buckling between the feed mechanism and the melt front location.

The cost efficiency, accessibility and simplicity of conventional FFF printers are key to their success. This is achieved by careful selection of components and technologies [30]. The use of open-loop stepper motors for actuation and the concept that the material feed mechanism functions both as the filament feeder and extrusion pressure generator are examples of this efficiency. A RepRap printer goes even further by striving towards using components, which are all re-producible by the printer itself.

2.1.2 Electronic Components

It is critical that the electronics of a FFF printer are open source and affordable, in order to maintain the cost efficiency, accessibility and simplicity of this type of manufacturing process. The electronics of these printers are therefore based on the well-known open-source electronics platform, called Arduino, which is an easy-to-use hardware and software platform. The first Arduino boards used 8-bit microcontrollers. This has since grown to vast variety of products, sensor modules and controller boards [37].

The FFF printer electronic functional blocks, as derived from [38, 39], are shown in Fig. 2.2. The firmware and electronics convert the electric power, control commands and sensor measurement input into the physical actuation (i.e. the machine output) required to print a part.

All component blocks interface with the main microcontroller, while the machine requires three communication interfaces with the outside world, for (1) programming the controller, (2) receiving commands for a host PC or command file and (3) for real-time user interaction. The communication interface with a host PC is mostly controlled with an Atmega 16-U2. Other solutions include direct USB connection to the microcontroller or a wireless interface. The user can also interface with the printer directly through a LCD interface with e.g. a touch screen or push buttons.

The firmware, which is the software program flashed onto the microcontroller, controls the printer and is introduced in Section 2.5. It is loaded through a programming interface port, for example a Joint Test Action Group (JTAG) connection or a USB port. Arduino based printers can be programmed with a USB connection, which emulates a serial port, with an Atmega 16-U2 microcontroller. On board memory is used to store configuration data. This means that the firmware functioning can be modified without re-flashing the microcontroller by only changing the configuration parameters.

The power management should supply both small signal electronics and higher power as required by the heaters, motors and fans.

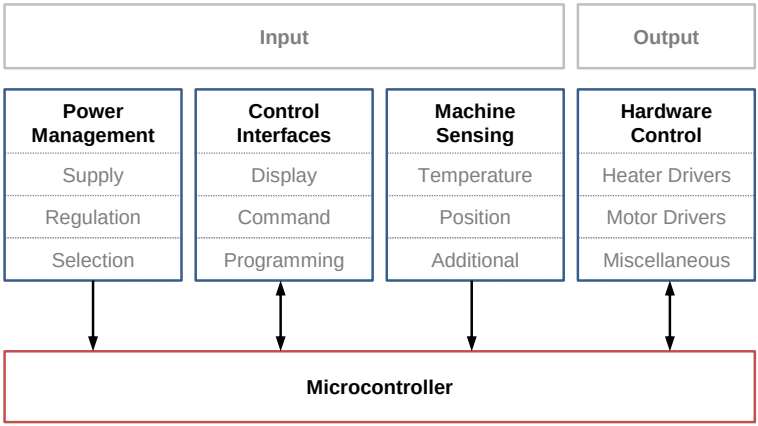


Figure 2.2: Electronic components of an FFF printer, grouped with the input and output functional categories. All component blocks interface with the main microcontroller, as indicated by the directions of the arrows.

The microcontroller also senses the machine state and environment with an array of sensors. The basic sensor set includes temperature sensors and travel end-stop locations (limit switches). Additionally, some printers also have bed level and filament presence sensors. Basic position sensing is achieved, however, with limit switches. This is used to home the printer axes and to prevent tool crashes at the end of travel. The temperature is measured with thermocouples, thermistors or Resistance Temperature Detectors (RTDs) (see Section 2.9 for an overview of FFF printer monitoring and closed-loop control).

The machine output is generated by driving the heaters and motors, which requires power electronics to convert the digital, small signal, control inputs to high current control outputs. Printers also use fans and other devices, for example, a case locking mechanism or printer illumination lights. All of these devices require a suitable DC power supply and a method to control the supplied voltage, for example to control the heater temperature with Pulse Width Modulation (PWM).

2.1.3 Auxiliary Components

General peripheral printer components are listed in Table 2.1, as derived from [31, 40].

The heated bed is essential to reduce part warping for printing with acrylonitrile butadience styrene (ABS) and it is useful when printing with PLA. A heated build platform can also aid first layer adhesion.

Table 2.1: Auxiliary components of an FFF printer.

Component	Function
Heated Bed	Reduce part warping and increase first layer bed adhesion
Hot-End Fan	Improve overhang and bridge structures
Cold-End Fan	Improve heat break effectiveness
Locked Enclosure	Reduce safety risk
Heated Enclosure	Reduce part warping
Filament sensor	Detect filament presence or jam

Part warping occurs when thermal stress warps a printed object, which is caused by the rapid cooling of the deposited tracks. Printing a *raft* around the perimeter of the first layer can also reduce warping. Additionally, a heated enclosure around the whole build volume can be used to further address this issue. Such an enclosure also reduces the risk of fire, injury due to moving parts and printer particle emission, if fitted with an air filtering system [39, 41] (see also Section 2.7).

The hot-end fan is used to rapidly freeze a printed track. This is the exact opposite function of the heated bed and it is recommended to turn the fan off for the first layer(s). This fan, however, helps to form structures such as bridges or overhangs. These part features are built with tracks which do not have support material underneath it. The fan speed is controlled with PWM and it can be adjusted by G-code commands. Various slicer settings are used to define the fan speed for different object features [42].

The basic extruder design is a compromise between a long and short liquefier and heat break. A longer liquefier and heat break allows for fast printing, but requires more space, while a shorter extruder is more compact (see Section 2.8.13). The cold-end fan helps in this case as it improves the effectiveness of the cold break [43]. Furthermore, some extruder designs use water cooling instead of forced air convention to cool, increasing the heat break efficiency.

Larger prints can take a long time (for e.g 24 hours). This requires a robust, reproducible process, without increasing the cost of the machine significantly. Various printer state, filament presence and feed jam sensors are therefore available or proposed (see Section 2.9).

2.2 Multi Material or Multi Extrusion Fabrication

Fused Deposition of Multi Materials (FDMM) [44] is also possible with FFF printers and it is one of the key advantages of this process type [16].

Example applications and methods of multi-material printing are listed in Table 2.2.

Multi-material FFF can be realised by using a mixing nozzle, more than one extruder or a hybrid process. Care needs to be taken to reduce unwanted extrusion during material change, i.e. a sharp contrast between materials can be challenging [39]. Another technique, which is not necessarily *multi-material*, is to vary the extrusion temperature or speed to produce a visual or material property variation.

The possibilities with multi material printing are exciting and most of these techniques rely on the basic extruder concept. This necessitates a deeper understanding of the FFF extrusion process, as discussed in the remainder of this chapter.

2.3 Drive Train Configuration

The RF1000 printer, sold by Conrad Electronics [33] in 2015, was used in this work and is shown in Fig. 2.3. The ceramic (white) heated build plate sits on the Y-axis carriage, which is displaced in Z-axis with two ball screws. The extruder is mounted on the X-axis carriage. Both the X- and Y-axes use a pulley and belt to drive a carriage, which rides on a lubricated linear bearing guide rail. The Z-axis is driven by belt drive, which actuates the two ball screws, which have recirculating ball bearings. The linear bearings also have caged ball bearings which reduces noise and generates a smooth motion. The bed is 235 mm long (Y-axis) and 250 mm wide (X-axis), while the maximum Z-displacement is 200 mm.

2.3.1 Parallel and Cartesian Configurations

FFF printers are arranged as Cartesian, parallel manipulators or a combinations of the two. The RF1000 uses a *serial Cartesian X-Head, YZ-bed* drive train configuration. Parallel manipulators are, for example, the *delta* and *coreXY* configurations, where a delta configuration can be realised with a Steward Platform layout or a linear delta robot configuration (an example of this is the *Rostock* printer). A coreXY configuration positions the head using an XY parallel manipulator, while the Z-axis moves the bed. It is therefore a *partially parallel Cartesian* drive train. SCARA (Selective Compliance Articulated Robot Arm) and polar configurations are also used, but are not as popular.

Parallel manipulators have the advantage that the motors can be stationary, reducing the load and therefore increasing the possible movement speed and acceleration, but this also requires additional calculations to convert the G-code Cartesian coordinates to machine coordinates.

The RF1000 printer addresses the motor weight issue by placing the extruder on the X-axis. The Z-axis, which does not require fast acceleration, as compared to the X and Y-axes, moves the bed as well as the entire Y-axis. This, however, creates a possible bed stability issue, since the bed is balanced on top of the Y-axis. Any roll or pitch (rotation around the Y or X-axes) will result in an unwanted Z-displacement of the bed.

Table 2.2: Applications of multi-material printing.

Application	**Description**	**Ref.**
Soluble supports	Printing of an additional soluble material for support structures	[28]
Multi-colour printing	Different colours of filament are used for increased object colour resolution	[39]
Multi-Functionality	For example, conductive and non-conductive filament	[41]
Increasing Print Speed	The same material is printed by two extruders concurrently, or a large and a small nozzle is used. The low resolution features, e.g. part infill, are printed fast and high resolution features slower, with the smaller diameter nozzle	[45]
Material Embedding	A composite fibre or wire is inserted into the track, for improved part strength or to introduce a new functionality	[46, 47]
Particle Embedding	Micro or nano particles are embedded into the polymer matrix to achieve special functionality. For example, magnetic properties	[48, 49]
Component Embedding	A part, for example, a nut or bolt, is inserted mid-print, embedding it in the object	
Track Painting	The extruded bead can be *painted* for visual effect or to add new functionality, for example, using an inkjet to print conductive silver nano-particle ink on the part	[50, 51]

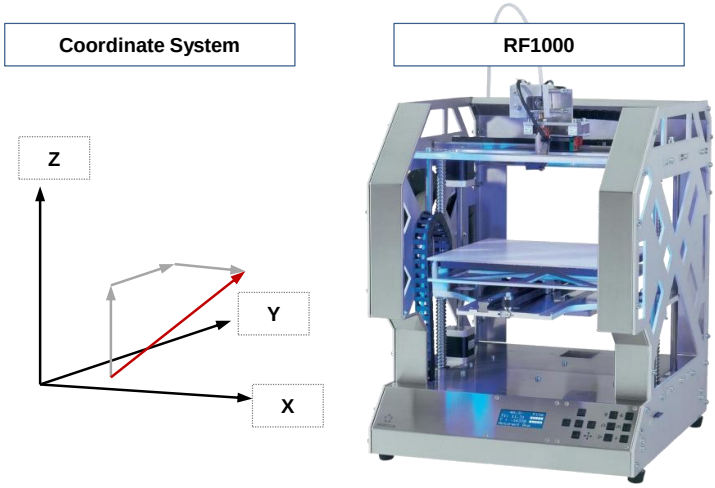


Figure 2.3: RF1000 (left) coordinate system and (right) mechanics. The red line indicated the position vector from the object origin to the nozzle [33].

2.3.2 Build Platform Z-Error

A photo of the Y-axis and the print bed is shown in Fig. 2.4. The ceramic build bed with the attached heater rests on four support pins, which are screwed into the mounting plates, fixed onto the Y-carriage. The plates are made from aluminium to reduce the mass, which needs to be moved. The entire Y-axis, including the rail and stepper motor, is mounted on the Z-carriage.

A sequential Cartesian configuration has six positioning errors per axis as well as three squareness errors, which results in 21 rigid body errors. The six errors are the linear displacement, the straightness (vertical and horizontal) and the angular (roll, pitch and yaw) errors. The THK SSR15-XW linear rails used by the RF1000 specifies a (0.02 ± 0.07) mm height tolerance and a (0.020 ± 0.006) mm width tolerance, for a normal grade rail [52].

A maximum bed level Z-error, z_e, can be estimated with $z_e = w_b \sin\theta$. The angle θ is equal to $\tan(zr/wr)$, where zr is the maximum rail height tolerance, w_b is bed width and wr the rail width.

A value for z_e of 0.233 mm is determined for the RF1000 printer with w_b, wr and zr equal to 250 mm, 15 mm and 0.015 mm. Note that z_e is almost one standard, low resolution, layer height (0.3 mm). The smallest standard layer height for this printer is 0.1 mm.

The maximum static deformation of the build plate $wmax$ will be at the plate centre, if it is simply supported and a uniform load p is applied. The deformation can be approximated with Eq. (2.1), if the plate is square

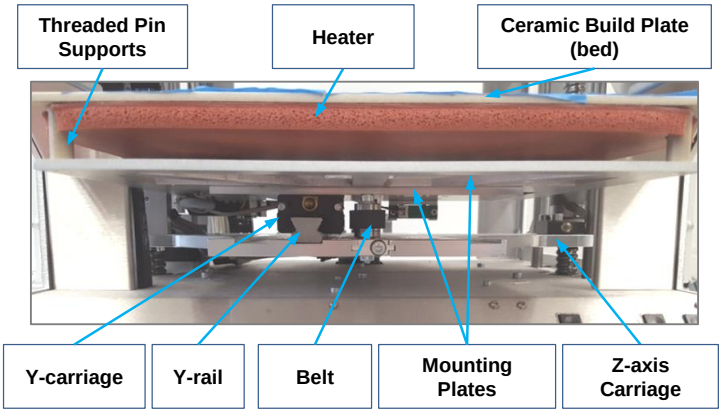


Figure 2.4: RF1000 Y-axis and build platform mounting.

and relatively thin, where D is the plate rigidity, E the Young's modulus, h the plate thickness, $v_{poisson}$ Poisson's ratio and a the length of the square plate [53].

$$D = \frac{Eh^3}{12\left(1 - v_{poisson}^2\right)} \tag{2.1a}$$

$$w_{max} \approx \frac{4a^4 p}{\pi^6 D} \tag{2.1b}$$

A maximum deflection of 23 µm is estimated for the RF1000 plate, assuming a square plate with length and height of 250 mm and 4 mm. The values for E and $v_{poisson}$ were taken as 21 GPa and 0.07, while a mass of 1 kg was assumed to be evenly distributed over the plate surface.

2.4 The Stepper Motor Motion Tool Chain

The positioning accuracy of the axes, including the feed mechanism axis, is a fundamental contributor to final part accuracy. Each axis must be accurately repositioned according to the G-code command. This is realised by the G-code interpreter (see Section 2.5), which accelerates and decelerates each axis so that the required position can be reached. It is, therefore, important to consider the relationship between the generated

electronic signals and the mechanical actuation, which is realised through the electric motors.

The RF1000 and many of the desktop 3D printers use stepper motors, as motivated by the RepRap design principles of cost effectiveness and accessibility. Indeed, it is almost standardised that 4 wire, bi-polar, hybrid stepper motors are used in these printers. These motors are discussed next.

Stepper motors do not require complex control electronics or brushes, which wear out. They can achieve accurate, reliable, low cost positioning and can generate high torque at low speeds. Disadvantages of stepper motors are, however, possibility of missed steps (in open loop control), limited torque at high speeds, step ringing and the need for microstepping to increase positioning resolution [54].

This section first reviews the drive chain hardware components. The limitations and requirements for the stepper motor torque, positioning and speed are considered next as well as the stepper motor driver influence on the printer positioning system.

The components used to move the printer axes, as required by the G-code commands, are shown in Fig. 2.5. The interpreter is responsible for parsing and executing the G-code, which specifies the axes, position (end-point) and target speed. This is done by building the velocity trajectory and then generating the required direction and step pulse-train in real-time, using the parsed G-code and the machine model (see Section 2.6). The step direction and pulse-train digital signals are sent to a motor driver IC, which converts the low current inputs into a high current output. The motor driver also manages the phase sequence timing. Finally, the stepper motor converts the electrical energy into the mechanical motion.

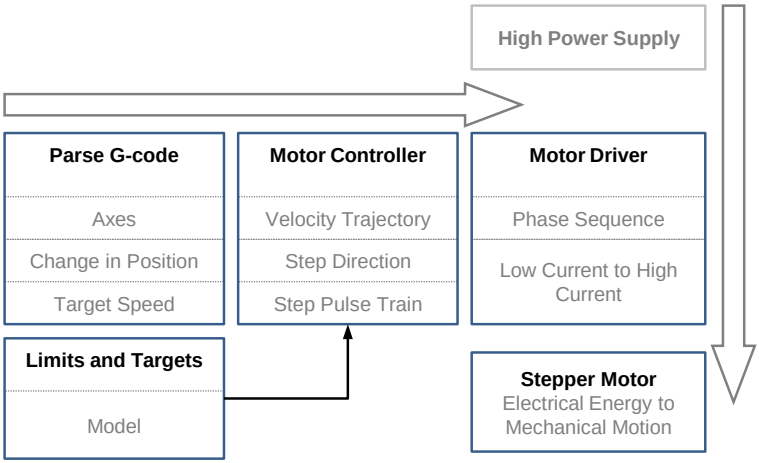


Figure 2.5: Motion generation tool chain.

2.4.1 Stepper Motor Torque Curve

The RF1000 uses four *Wantai* 42BYGH-W811 stepper motors, which have holding torque of 4800 g · cm (approximately 0.48 N · m), a stepping precision of 5 % and a step angle of 1.8°, which means there is 200 full steps per revolution [55]. The stepping precision does not accumulate and it is only a measure of the positioning accuracy of each full step.

It is difficult to obtain detailed information regarding these motors as datasheets usually only report 24 V torque curves, while many RepRap stepper motors operate at 12 V. A maker community contributor, therefore, measured the torque curve for few popular stepper motors, with a custom built torque rig and a RepRap motor driver (Pololu A4988) at 12 V [56]. The results are shown in Fig. 2.6.

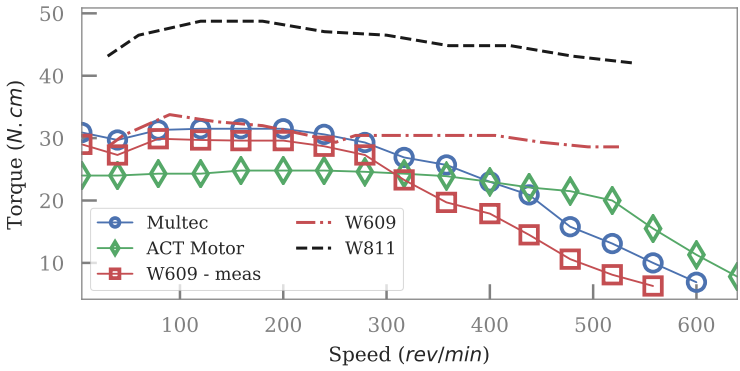


Figure 2.6: Measured torque versus rotation speed, from [56], with unfilled circles, squares and diamonds for the Multec, ACT Motor and the Wantai 42BYGH-W609 stepper motors. The dash-dot line is from the W609 datasheet and the dashed line for the RF1000 W811 stepper motor [55].

The W609 motor is similar to the W881 motors used by RF1000 printer, but has a lower holding-torque (4000 g · cm) [55]. The measured data show that the lower inductance motors with higher rated current are able to generate more torque at faster speeds. Note that the measured data does not represent a motor test result, but is intended only as a design guide [56].

The torque values claimed by the datasheets for both the W609 and W811 stepper motors are also shown. The torque curves supplied by datasheets indicate a slower torque decrease with speed increase, than the test results. This decrease becomes significant at approximately 300 rpm.

2.4.2 Stepper Motor Microstepping Torque

Intermediary steps between the full and half steps, called microsteps, are used to reduce the ringing effect at low speeds and noise at higher speeds. Microsteps also reduce resonance issues. It is created by applying a step like sinusoidal current curve to the motor windings, where the individual coil currents are 90° out of phase, for a two phase motor. Note that microstepping does not necessarily create higher precision motion, due to static friction and linearity problems [57].

Static friction increases the minimum torque required to rotate the rotor. This means that the absolute value of the ideal sinusoidal torque curve, which is less than the friction torque, is clipped to zero. This creates a dead zone or band. That is, the commanded microstep (change in coil current) does not result in any rotor rotation until the generated torque is higher than the static friction limit. This also creates the ringing effect or results in missed steps [58]. Backlash, positional hysteresis, can also occur when the motor changes direction, due to this dead band, since several steps are required before the shaft rotation direction changes.

Another limiting effect is the non-sinusoidal torque output of a stepper motor, when supplied with an ideal sinusoidal current. This is due to the motor construction and design. For example, the design of a motor can be optimised for full stepping and microstepping such a motor will lead to missed steps. The stepper motor is also driven by a quantized current, with a resolution depending on the number of microsteps. The resulting quantization error can be reduced by either selecting optimal shaft positions or by increasing the allowable output torque variance, i.e. reduce the required torque to move the stage [57].

Nevertheless, microstepping is used to increase the step resolution and to create a smoother motion, by reducing vibration. A 1.8° stepper motor with 200 full steps per revolution, will ideally be able to achieve 51 200 steps with a 256 microstep resolution. The achievable resolution is, however, limited by the decreasing incremental torque with increasing number of microsteps.

The incremental torque increase, τ_μ, for an ideal stepper motor can be calculated with Eq. (2.2), where τ_h is the full step holding torque and S_μ the number of microsteps [59].

$$\tau_\mu = \tau_h \sin\left(\frac{\pi}{2S_\mu}\right) \tag{2.2}$$

The percentage incremental torque decrease calculated with Eq. (2.2) is given in Table 2.3. At 16 microsteps the incremental torque is already decreased to less than 10 % of the full step holding torque.

The required pulses per second (PPS) to reach a speed of 50 mm/s are listed in Table 2.3. This is calculated with Eq. (2.3) and is the number of step pulses per second f_v that both the controller and driver must handle,

Table 2.3: Decrease in incremental torque [59] and increasing step pulses per second (PPS), due to increasing the number of microsteps for the RF1000 X-axis, for a speed of 50 mm/s

Microsteps	**Incremental Torque**	**PPS**
S_μ	%	kHz
1	100.0	0.2
2	70.7	0.5
4	38.3	1.0
8	19.5	1.9
16	9.8	3.8
32	4.9	7.6

where v is the speed, n_{spr} the number of full steps per millimetre and S_μ the number of microsteps per full step.

$$f_v = v n_{spr} S_\mu \tag{2.3}$$

The RF1000 printer should be able to achieve a maximum step frequency of 40 kHz, with firmware based on the Repetier-Firmware distribution and an Arduino based AVR microcontroller board [60]. This is a little more than 128 microsteps per full step.

2.4.3 The Millimetre to Steps Conversion Factor

The total distance is set by the G-code command in millimetre. The firmware, however, requires the relationship between the step pulse, stepper motor shaft position and the axis displacement. This is determined with a constant conversion factor S_{cnv} in steps/mm, which is calculated with Eq. (2.4) for the RF1000 printer X and Y axes. The calculations for the Z and E-axes are slightly different, but use the same principle.

$$C_{pulley} = p_{belt} N_{pulley} \tag{2.4a}$$

$$S_{cnv} = \frac{n_{spr} S_\mu}{C_{pulley}} \tag{2.4b}$$

$$n_{total} = S_{cnv} L_{axis} \tag{2.4c}$$

The stepper motor shaft drives the carriage on a linear bearing with a pulley and belt drive. The RF1000 belt has a pitch, p_{belt}, of 2.1 mm and a pulley with 20 teeth (N_{pulley}) (see [61] for the RF1000 printer details and [62] for RepRap printers).

The total number of microsteps n_{total} required to move an axis a length of L_{axis} is determined with Eq. (2.4c), noting that the RF1000 printer uses 32 microsteps per full step and has 1.8° step angle motors, with $n_{spr} = 360/1.8 = 200$ full steps per revolution.

The conversion factors are stored in the printer EEPROM or hard coded into the firmware. The values for the RF1000 printer are given in Table 2.4.

Table 2.4: The steps to millimetre conversion factors for the RF1000 printer, as stored in the on-board memory.

Axis	**Conversion Factor** steps/mm	**Minimum Step Size** µm
X and Y	$4.761875 S_\mu$	6.6
Z	$80.000000 S_\mu$	0.4
E	$8.750000 S_\mu$	3.6

The theoretical minimum step size for each axis can be calculated by inverting S_{cnv}. The actual step resolution will, however, be limited by the decreasing incremental torque. This is further complicated by the use of *Quad Stepping*, which is used to smooth the velocity trajectory and reduces the actual number of microsteps by factor of four (see Section 2.6).

The actual steps per millimetre factors for each axis can be determined with a short calibration routine, by determining the difference between the commanded and measured position change (see [39], for an example).

The effective pitch diameter D_{feed} of the extruder gear, which is mounted on the feed mechanism motor, can be determined from the number of steps per millimetre for the extruder $S_{cnv,E}$, if this diameter is unknown. Note that this will be less than the external diameter of the gear and will depend on the pinch force used to grip the filament as well as the gear tooth design. The RF1000 feed gear pitch diameter is 7.276 mm, with Eq. (2.5), by reordering Eq. (2.4).

$$D_{feed} = \frac{n_{spr} S_\mu}{\pi S_{cnv,E}} \tag{2.5}$$

2.4.4 Required Motor Torque

The torque required to move an axis versus the torque generated by the stepper motor will determine if a position is reached or if steps are skipped, if an open loop controller is used. A basic model for determining the required torque is shown in Fig. 2.7 and calculated with Eq. (2.6) [63–65].

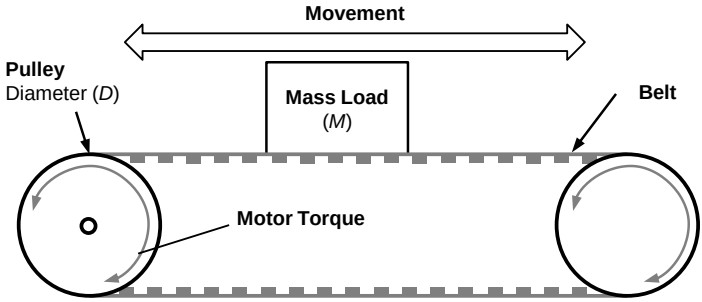


Figure 2.7: Model for estimating the torque required for the X- or Y-axis. The load is moved with a belt, which runs around two pulleys

The resultant torque required to accelerate (a) the load and the drive train is determined with Eq. (2.6c), where both the static and dynamic components are seen. The static component is the friction, which the motor must overcome and is a function of the gravitational acceleration g, pulley diameter D and mass m_L, as well as the friction coefficient μ and drive efficiency η. Note that τ_t is in $\mathrm{N \cdot m}$, J_T in $\mathrm{kg \cdot m^2}$ and mass in kg.

The total system inertia J_T is determined with Eq. (2.6a), which includes the mass of the load, belt and pulleys (m_L,m_B and m_P) as well as the motor inertia, J_M.

The acceleration a is converted from $\mathrm{rev/min^2}$ to $\mathrm{rad/s^2}$ with Eq. (2.6b). This is based on the assumption that the motor can achieve a certain speed v_S in step/s instantaneously, within a maximum lag of two full steps, which give a time lag Δt of $2/v_S$. The acceleration is determined

with $a_S = \Delta v/\Delta t = \Delta v^2/2$ (in step/s^2).

$$J_T = \frac{1}{4}\left(m_L + m_B + m_P\right) D^2 + J_M \tag{2.6a}$$

$$a = \pi n_{spr}\left(\frac{v_{rpm}}{60}\right)^2 \tag{2.6b}$$

$$\tau_t = \frac{m_L g D \mu}{2\eta} + J_T a \tag{2.6c}$$

The estimated torque required for the RF1000 X-axis, which carries the extruder, is shown in Fig. 2.8. The following values were used to perform the calculation. A total mass of 0.760 kg for the load is assumed, based on a mass of 340 g, 190 g, 200 g, 20 g and 10 g for the motor, carriage, feed mechanism, mounting bracket and liquefier, while the mass of the belt and pulley are assumed to be 10 g and 7 g. The motor inertia is 68 gcm^2 according to the datasheet. The pulley diameter D is estimated with C_{pulley}/π. The friction and efficiency coefficients are assumed to be 0.3 and 0.85.

Figure 2.8 indicates the speed in three different units, namely: rpm, mm/s and PPS, with unit conversion based on the RF1000 configuration. The revolutions per minute are important for the motor characteristics, the millimetres per second axis for the printing process as well as the force dynamics, while the pulses per seconds are important for the processor, control algorithm and motor driver design.

The torque versus speed, according to the motor datasheet, is also shown. The upper dashed horizontal line is the average maximum torque, which the motor should be able to generate. Microstepping, however, reduces the average torque with about 71 % [66]. An additional safety factor of two is used to further reduce the estimated maximum motor torque. This is essential since the motor will be used in open loop control and step loss must therefore be minimised. The resulting estimated maximum torque is depicted with the solid horizontal line.

The reduced maximum torque intersects the predicted required torque at about 150 rpm or 100 mm/s. Interestingly, the standard slicer parameters also set the maximum printer speed to 100 mm/s.

The lower dashed horizontal line is the estimated friction torque or the minimum torque required to move the X-axis from rest. The two short solid vertical lines indicate the maximum motion joining speed (left) and maximum start speed (right), as set in the RF1000 configuration.

The maximum motion joining speed is 10 mm/s for the X and Y-axes and is used to determine the deceleration required between two print lines, if they are in different directions. This is discussed further in Section 2.6.1.

The maximum starting speed for the extruder motor is set as 18 mm/s. This is the fastest instantaneous speed, which a motor can achieve without acceleration. A maximum starting stepping frequency f_S in steps/s is

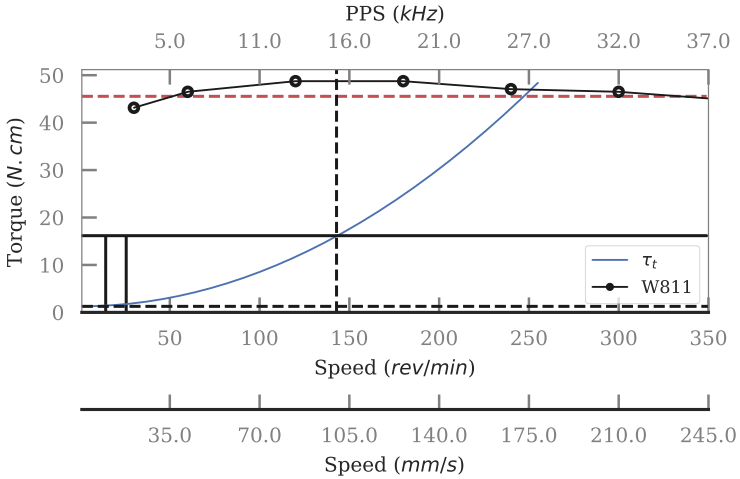


Figure 2.8: The W811 motor and estimated required torque (τ_t) for the printer X-axis versus speed. See text for explanation of the horizontal and vertical lines.

defined by some datasheets, but is unfortunately not specified in the W811 or W609 datasheets.

2.4.5 The Stepper Motor Driver

An important part of the motion driven chain is the motor driver. This IC must convert the step pulse train and direction signals into a quantised sinusoidal current versus time curve, to create smooth and accurate motion. The accuracy of the generated current level for each step is important since the motor torque (and therefore speed and positioning) is proportional to the current. The ideal instantaneous current $I(t)$ in Eq. (2.7) is determined by the applied voltage V, winding resistance R, coil inductance L and the electrical time constant $\tau = \frac{L}{R}$.

$$I(t) = \frac{V}{R}\left(1 - e^{-\frac{t}{\tau}}\right) \tag{2.7}$$

The time constant limits the maximum motor speed, since the current in the winding increases slowly, relative to fast step pulsing frequencies, if a constant voltage is applied. This limits the maximum current reached

and therefore limits the torque and speed of the motor [58]. This can be improved with an $L/4R$ control scheme, by introducing a series resistance to the coil and by increasing the applied voltage. Such a scheme is, however, inefficient, since the extra resistance wastes power with $P = I^2R$.

A constant current driver is therefore preferred, which applies a voltage higher than the rated voltage and then *chops* the current as soon as it is more than the set maximum winding current. A popular technique to create such a circuit is to use PWM (Pulse Width Modulation) with current feedback. Current chopping solves the time constant problem and reduces power dissipation. A chopper driver IC generates more torque at slower speeds and can also achieve faster top speeds, than a $L/4R$ controller [67].

The current rise time t_{cc} can be used to estimate motor maximum speed, by solving t in Eq. (2.7) for given supply voltage, current limit and stepper motor characteristics.

The maximum speed (v_{max}) of the two Wantai motors, presented previously in Fig. 2.6, are compared using $v_{max} = 1000/(S_{full}tcc)$, a supply voltage of 12 V and the rated motor current. The results are given in Table 2.5.

Table 2.5: Maximum motor speed based on coil inductance.

	W609	**W811**	**Unit**
I	1.70	2.50	A
L	3.0	1.80	mH
R	2.00	1.25	Ω
t_{cc}	0.50	0.43	ms
v_{max}	420	483	mm/s

The default maximum speed for the RF1000 X- and Y-axis is 500 mm/s, which agrees with the estimated speed limit for the W811 stepper motor of 483 mm/s, which is also slightly faster than the estimated maximum speed for the W609 motor.

The accuracy of the motor speed will also depend on the motor driver accuracy, which depends on the current output accuracy of a PWM chopping circuit. This, in turn, depends on the voltage reference V_{ref}, sense resistor R_{sense} and the current ripple, since the output current level is proportional to V_{ref}/R_{sense} and the current variation to the ripple. The current ripple is therefore a function of the supply voltage, driver and motor properties, as well as the chopping frequency. Increasing the chopping frequency reduces the ripple, but is limited by thermal switching losses and noise [58, 68].

Listing 2.1: G-code syntax where "n" is numeric and "A" and "B" are alphabetical characters.

```
Gnnn␣;<a_comment_string>
Gnnn␣A␣B
G1␣Xnnn.nnn␣Ynnn.nnn␣Ennn.nnn␣Fnnn.nnn
Mnnn␣Snnn
```

2.5 The G-Code Interpreter

The main microcontroller runs a firmware program called the interpreter, which is introduced in this section, with the focus on how a G-code command is translated into physical motion.

2.5.1 G-Code Syntax

G-codes, which is short for *general command*, are not just commands, but actually a standardised programming language. The "*NIST RS274NGC Interpreter*", where NGC stands for Next Generation Controller, is explained by Kramer, Proctor, and Messina [36].

RepRap G-codes are based on this work, but do not fully adhere to the G-code standard [34]. Furthermore, different firmware branches implement certain aspects or commands differently. The syntax for a RepRap type G-code is shown in Listing 2.1

The first line shows the syntax for a single command, followed by a comment string. An example of such a command is **G28**, which homes all the axes or **G90**, which tells the interpreter to use absolute positioning. The second line specifies a list of variables to which the given command must be applied. An example of such a command is "**G28 X Y**", which specifies that only X and Y-axes must be moved to the home position. This can be expanded to include values, as shown in the third line, which shows the syntax for the important **G1** command discussed later. The last line shows the syntax for an *M* command. The syntax is the same as for a *G*-type command. An example of this is "**M106 S127**" which sets the fan speed to 127, which is approximately 50 % of the fan output power.

The standard defines *modal* and *non-modal* commands. A modal command changes the state of the machine and is active until the next modal command, while non-modal commands only affect the command lines on which they occur. Thirteen modal groups are defined and one group for non-modal commands. Group one is the *motion* group, which contains machine positioning commands. Apart from the G-codes, there are also M,T,F-codes: *miscellaneous*, *tool*, and *feedrate* commands (see [34, 36] for a definition and description of all commands).

The syntax allows for only one modal group of commands per line. For example, only one movement command. A line can, however, request the simultaneous actuation of several axes. The commands or fields are separated with white spaces and a line is ended with an end-of-line marker. Comments, checksums and line numbers are optional. Each field starts with a letter, for example *G*, and can be followed with a value. If a field does not have a value, it is considered to be a flag. A parameter value can be an integer, fraction or a quoted string [34, 36].

2.5.2 The G1 Command

The **G1** movement command, along with the feed speed, set with the *F* command, are of significance in this work. The **G1** command requests that the printer moves linearly to the destination points of the specified axes. At least one axis must be defined, e.g. "**G1 X200**", will move the X-axis at the current feed speed to coordinate position 200 mm (if the printer is set to use absolute positioning and millimetres).

An FFF printer has four axes, namely *X,Y,Z* and *E*, where the E-axis is the extruder (feed mechanism). Note that the feed rate (*F*) is always in millimetres per minute.

According to [36] the **G1** command should initiate a *coordinated linear motion*. This is a displacement of all the axes to the defined end points completed at the same time. The speed trajectory of each axis must therefore be adjusted so that all the axes move the same fraction of the required displacement. That is, all the axes are slowed down so that the movement is completed in the minimum movement time of the slowest axis.

The execution order of the commands in a single line is also important. In the standard, the feed speed setting command is executed first and then the movement command. The RepRap firmware notes that the feed speed will linearly accelerate, along with any other axis, which ensures that volumetric flow rate is synchronised with the XY positioning.

2.5.3 The Fused Filament Fabrication Interpreter

The firmware of RF1000 printer (version 1.10) is based on the *Repetier firmware* branch, which, in turn, was derived from *Marlin*. The *Marlin* firmware was forked from *Sprinter*, as well as *Grbl*, while the *Sprinter* firmware branch was essentially derived from *Generation2* and the original RepRap *FiveD Gcode* software repositories [69].

Most open-source 3D printers use some aspects *Grbl*, which is a high performance, Arduino compatible, G-code interpreter for controlling the motion of machines. It was developed by Simen Svale Skogsrud in 2009 and it is claimed that: "if the maker movement was an industry, Grbl would be the industry standard" [70].

All of these firmware variants must implement a G-code interpreter. A block diagram of the interpreter is shown in Fig. 2.9.

The interpreter is driven by a *driver* through a communication interface. The driver can be a host PC, the user or a file on an SD card, which

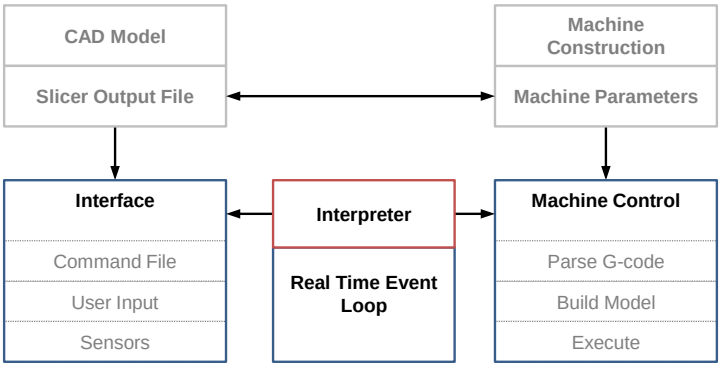


Figure 2.9: Tasks of the Interpreter

communicates through a flow control protocol. The commands themselves are normally generated by a slicer. The user or print server, however, can also supply commands in real-time, as well as request information from the machine, while the interpreter reads and stores the printer sensor outputs.

Ideally, all G-codes should work on any machine with a G-code interpreter, which serves as an abstraction layer, so that the driver (i.e. the slicer) does not need to know the intricacies of how to actuate a certain axis on a specific machine. The driver only needs to specify the outcome, or the end position, which must be achieved at a certain speed.

A 3D Printer not only needs to convert the G-codes into machine function calls, but must also control the execution of it. This is called an Enhanced Machine Controller (EMC) by [36] and follows a two-step process. First a command line is read and parsed into machine function calls. These calls are then performed according to a specified order of execution [36].

The interpreter, however, does not know the machine configuration (construction) and this must be predefined either through hard coded values or through parameters stored in on-board memory. Note that it is important that the machine commands generated by the interpreter are both physically realistic and satisfy the requirement of the G-code command, as accurately as possible. This is made possible by the interpreter based model of the machine. Two models therefore exist: the slicer (driver) model of the printer and the firmware (interpreter) model.

An example of this is the maximum acceleration. Exceeding this limit can result in missing steps. The interpreter therefore tries to reach the commanded speed, without exceeding the acceleration limit, within the given distance. This limit is, however, determined by both the hardware and speed ramp model. That is, the slicer sets the acceleration limit at the start of the G-code file and assumes this is executed perfectly. The

firmware, however, uses an approximation model to determine the speed ramp, which can result in instantaneous speed or acceleration changes. This can lead to a difference in commanded and actual machine output, reducing the quality of the printing process.

Methods to address the difference in this specific example are (1) improving the mechanics and drive to increase the maximum possible acceleration, (2) reduction of the acceleration limit value, (3) improvement of the slicer model with better path planning (reducing the required acceleration), or (4) implementing a more complex acceleration ramp in the firmware. This is only an example and the realisation of the positioning by the interpreter is discussed further in Section 2.6. The point here is that all three (machine, interpreter and slicer) need to be synchronised to achieve better print results.

The firmware also handles parameter setting, the communication interface, user inputs, PID temperature control and watch-dog safety considerations, amongst others. These are however not discussed here, since the focus is the execution of a print move.

2.6 The Real-Time Velocity Profile

The interpreter must perform trajectory planning in order to execute a movement command. Smooth, fast and high resolution positioning are desired, but the stepper motor torque curve, load dynamics and computationally efficient implementation of the planning algorithm are challenges.

The motor drive chain, as discussed in Section 2.4, found that the microcontroller must generate a pulse train and direction digital signals to move an axis. The positional quantisation effect, as result of finite steps, can be reduced with microstepping, but this requires faster stepping rates. Furthermore, the step frequency should increase over time, if acceleration is required, and the efficiency of this calculation is critical for smooth operation. The independent motion of all the axes must also be combined to ensure that the print head is moved in a coordinated linear manner, relative to the build platform and that volumetric flow rate is synchronised during the print move.

Two issues, which significantly effect a stepper motor stage and lead to step loss, are jitter and jerk. Jitter is vibration during movement and jerk is the time derivative of acceleration. The output of modern current drivers is near perfect and only limited by the motor construction [57]. Recent improvement work is in the area of software development, where various techniques are used to reduce motion jitter and jerk. These concepts are discussed next, with the aid of practical examples by simulating a set of G-code commands. First, the concept of motion joining is presented, then the velocity profile and step pulse generation techniques. Following this, the Bresenham line approximation and multi-axis motion smoothing are discussed. The section is concluded by presenting the RF1000 firmware implementation, in context of the command buffer and motion planning routine.

Listing 2.2: Example G-code commands for Trajectory planning.

```
G1 X1.000 Y0.750 E0.042 F900.000
G1 X0.200 Y1.250 E0.043 F1200.000
```

2.6.1 Motion Joining

A set of **G1** movement commands are shown in Listing 2.2 and it is used in this section to understand how FFF printers achieve trajectory planning.

It is assumed that the printer is warmed up and at rest, at coordinates $0 \times 0 \times 0$. The first command moves the printer to the absolute XY coordinates 1 mm and 0.750 mm, with a target speed of 900 mm/min (15 mm/s). The second command request that the axes are displaced to 0.2 mm and 1.25 mm at 1200 mm/min (20 mm/s). This will mean, for the RF1000 Cartesian configuration, that the X-axis will need to change direction (first move right and then left), while the Y-axis can continue in the same direction.

Motion in FFF printers is mostly achieved in a piecewise linear fashion, where each linear section is a G-code command. The interpreter builds a movement plan for each axis, which creates the coordinated linear motion. This means that the speed vectors must be considered, so that each axis can be moved over different lengths, with different speeds and accelerations, but still reach the final destination at the same time. The speed vectors of the example G-codes are shown in Fig. 2.10. The two black lines are the command speed vectors and the red line the vector sum.

A stepper motor can not change speed instantaneously. For the example G-code this means that the X-axis should slow down and then accelerate again in the opposite direction. Jerk must therefore be minimised to an acceptable level. The stage can break or lose steps if the jerk is too extreme.

The jerk can be reduced by bringing the axis to a complete halt, before changing direction. This will result in accurate point-to-point positioning, but can create issues with the volumetric flow rate and machine vibration as well as increase the printing time. The maximum achievable linear speed will also be reduced for short moves.

A method which reduces jerk, but increases the maximum cornering speed is therefore required. This can be done by: (1) trading accuracy for speed, (2) reducing the maximum speed or (3) reducing the joining speed [71].

Accuracy can be traded for speed by not passing through the joining point at all, by creating a larger turning radius. A faster speed can therefore be used, but this sacrifices accuracy. The slicer algorithm can implement such an approach during the generation of the G-codes. A CAD object with two walls joined by 90° edge can be completed in three

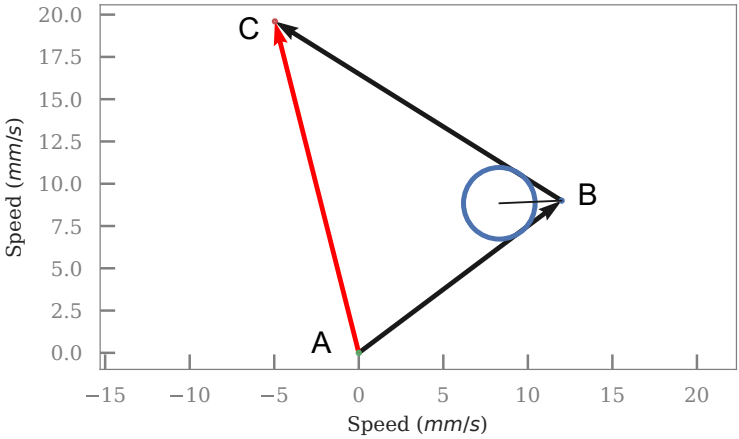


Figure 2.10: Calculation of the maximum joining speed. Black arrows is the commanded speed vectors, with a circle showing the maximum *velocity* turning radius.

moves instead, by inserting a small diagonal (45°) move between the two walls. Special, non-G1 commands, can also be used, if implemented in the firmware to facilitate such motion joining.

Stepper motors have a maximum starting speed, v_0, based on f_s, the maximum starting stepping frequency. In principle, any moves can be joined without changing the speed at all, if the total change in speed is less than v_s. This technique, however, introduces a lot of vibration and reduces the total printing speed significantly [71]. The maximum starting torque can be seen by comparing the friction torque with the acceleration torque required to move the axis in Fig. 2.8.

Motion joining can be improved by inspecting the next move. If the next movement requires the same velocity, then the deceleration is no longer required, which increases print speed and reduces vibrations. A deceleration calculation is required if the next move requires a slower speed or if the movement is in a different direction. The question is what is the maximum *safe* speed with which the printer can corner.

The Repetier firmware uses what it calls a *maximum jerk* value. It is clearly stated in the documentation that this jerk value is not the time derivative of acceleration [72]. A better word is possibly the maximum joining speed. The calculation is, however, straightforward. The maximum joining speed is calculated with Eq. (2.8), where the magnitude of the speed change Δv between speed vectors ($\vec{V_{AB}}$ and $\vec{V_{BC}}$) is limited

to be less than the maximum joining speed v_0.

$$\Delta v = \|\vec{V_{AB}} - \vec{V_{BC}}\| < v_0 \tag{2.8}$$

The implementation of this calculation is more complex. All the moves, with pre-calculated trajectory parameters, are stored in a buffer. The maximum speed, which can be reached during the move depends on the acceleration, deceleration and the distance. The deceleration distance, however, depends on the end speed of the move.

The movement joining algorithm therefore requires a reverse and a forward computation. The reverse calculation is performed first, from the last entry in the buffer to the third entry, where the first and second entry will most likely be executed soon or are currently being executed. The reverse calculation maximises the previous move end speed, so that Δv between the previous move-end and current move-start does not exceed v_0. A forward planner then determines the maximum speed, which can be reached (given the target maximum speed, distance and end speed). A buffered move is locked once both an optimal starting and end-speed are determined.

The firmware also estimates the time available to perform these calculations and can limit the number of iterations or reduce the movement speed, so that the calculations do not exceed the available processor time.

The maximum joining speed calculation (Eq. (2.8)) is computationally efficient, but uses a linear fit for a non-linear acceleration problem. This approximation only works if the v_0 limit is finely tuned for each printer. Furthermore, it can result in *choppy* or slow movement through sharp corners [73].

The Repetier-firmware addresses this issue by including a factor calculation, given in Eq. (2.9a), which scales the maximum speed of the previous vector with the *jerk* value. The joining speed v_{max} is then set as the minimum between the scaled entry speed and the non-scaled exit speed, with Eq. (2.9b) [72].

$$f = \min\left(1, v_0/\Delta v\right) \tag{2.9a}$$

$$v_{max} = \min\left(f\|\vec{V_{AB}}\|, \|\vec{V_{BC}}\|\right) \tag{2.9b}$$

The Grbl firmware addresses the motion joining issue by considering the virtual centripetal motion, assuming the tangential velocity is zero and by finding the cornering radius, R, at a virtual offset δ to determine the maximum speed through the corner (see Fig. 2.10). Such an approach is more accurate for all ranges of motion and can be computed with only two square roots [73].

The circle in Fig. 2.10 represents the cornering radius. A short solid line connects the circle centre point with the joining point of the two vectors and the angle between the two vectors is θ. The right triangle formed by

the circle centre, perpendicular radius length to the vector and the joining point can be used to calculate R with Eq. (2.10a) and θ with Eq. (2.10b). The required cosine and the inverse cosine calculations can conveniently be replaced with the half-angle identity $\sin(\theta/2) = \pm\sqrt{\frac{1-\cos\theta}{2}}$.

$$R = \delta \frac{\sin(\theta/2)}{1 - \sin(\theta/2)} \tag{2.10a}$$

$$\cos(\theta) = \frac{\vec{V_{AB}} \cdot \vec{V_{BC}}}{\|\vec{V_{AB}}\|\|\vec{V_{BC}}\|} \tag{2.10b}$$

The calculated normal junction speed v_j is calculated with Eq. (2.11) at the predefined maximum acceleration a_{max}. The maximum junction speed v_{max} is the minimum cornering speed, which is either the minimum speed of the two speed vectors or the calculated junction speed v_j.

$$v_j = \sqrt{a_{max} R} \tag{2.11a}$$

$$v_{max} = \min\left(v_j, \vec{V_{AB}}, \vec{V_{BC}}\right) \tag{2.11b}$$

The three different joining speed calculations are compared in Fig. 2.11. The example G-codes (Listing 2.2) were used. A range of corners were simulated by rotating the second vector from 0 – 179°, where an angle of 0° means that there is no change in direction. The plot on the left (A) shows the vectors for selected angles (every 30°) and (B) shows the calculated maximum joining speed with an angle increment of 1°.

The *constrained* method is based on Eq. (2.8), the *factor* method is determined with Eq. (2.9b) and the *normal* method with Eq. (2.11). The lower dashed horizontal line indicates the RF1000 printer configuration *maxJerk* parameter. Standard values from the Grbl firmware was used for the *normal* line calculation, namely 0.01 mm for the virtual offset δ. The maximum acceleration a_{max} is taken from RF1000 setting for the X and Y-axes as $1000\,\mathrm{mm/s^2}$.

This graph motivates the use of either Eq. (2.9b) or Eq. (2.11b) for the calculation of v_{max}, since the constraining method does not reduce speed proportional to the angle between the vectors. The Gbrl method, however, seems to work better than the RF1000 technique, since it gradually reduces the joining speed to zero in sharp corners.

2.6.2 The Velocity Trajectory

The velocity trajectory for each axis needs to be planned before execution. This is the speed versus time profile converted to the direction and a

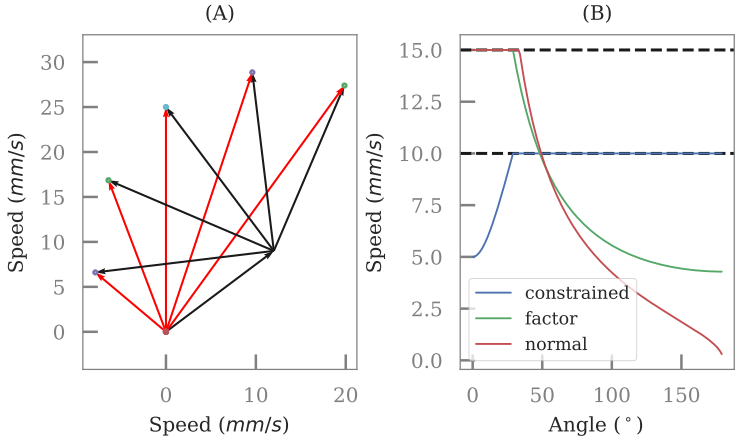


Figure 2.11: Comparison of different motion joining techniques. (A) shows the generated vectors with an angle increment of 30° and (B) is the resulting joining speed for the three different methods discussed in the text.

step pulse train digital outputs, which is converted to the microstepped sinusoidal current curve for each motor coil, which produces the stage displacement. The digital step pulse train is separated with a certain delay time. This delay time determines the speed and needs to be calculated in real time by the firmware, given the current speed and desired acceleration or deceleration. The velocity trajectory, however, needs to be determined before the movement execution.

Various profile orders are used to build this trajectory. First-order planning was used by the original RepRap printers. It is the most straightforward approach and uses only a constant speed profile. This limits the maximum achievable speed, due to the maximum starting frequency of the stepper and vibration issues.

Second-order profiles introduce constant acceleration and deceleration. These profiles are used by the Grbl and Repetier firmware. The calculation of the step pulse train to achieve the required acceleration is done in different ways. The first approach simply reduces the step period with an arbitrary value until the desired speed is reached. The next method calculates the required step delay, at each step. This is more accurate, but requires more computational power. The calculation requirement can be reduced by performing it at a constant frequency, slower than the stepping frequency. This reduces the velocity accuracy, but the error can be negligible if the calculations are performed fast enough [74, 75]. There are therefore two options defining *when* to calculate the next time delay

in real time for second-order profiles. It can either be done on a step by step basis or with a constant update frequency.

Higher order models are also used. A fifth order trajectory profile for a linear delta robot configuration FFF printer is presented in [76]. The *TinyG* and *G2 Core* multi-axis motion control systems use a 6th order, jerk-controlled, real-time planner. The main reason for implementing such is to reduce the occurrence of infinite jerks, which is a step change in acceleration, to better simulate the physical capabilities of the positing system. This results in less wear on parts, smoother operation, reduced vibrations and higher speeds [77].

The firmware generation of the second order velocity profiles are investigated here, by simulating the example G-codes with a program written with the Python programming language. The velocity and distance moved for each line are shown in Fig. 2.12. In the simulation only the X and Y-axes are considered. The E-axis is included in the same way as the X and Y-axes, but the Z-axis is treated a little differently, since it is only stepped once, after layer completion.

The use of constant acceleration results in two sets of semi-symmetric trapezoidal speed profiles, with different start and end speeds for the example G-codes. The displacement is non-linear during the acceleration and deceleration phases of each move and linear during the *coasting* section of the move. The trajectory will become triangular, if the target speed cannot be reached within the time span.

A second order speed trajectory is defined by the start-, end- and target-speed, as well as the total distance. Note that the **G1** command only supplies the target speed and distance. The start and end speeds are determined with a motion joining technique. The deceleration and the acceleration distances (or time points) are also important. These are used for the real-time calculation of the step frequency, in a computationally efficient manner, as presented later.

The ideal velocity trajectory used here is determined as follows. The maximum change in speed Δv_i for the start Δvs and the end Δve section are determined with $\Delta v_i = v_t - v_i$. The distances moved during the ramp sections (S_i) are determined with Eq. (2.12), by assuming a maximum acceleration $amax$ for each axis.

$$S_i = \Delta v_i^2 / amax \tag{2.12}$$

Note that the trajectory profile will change to a triangular profile if $S_{start} + S_{end} >= \Delta S$, where ΔS is the total G-code commanded distance. The time required for each section in the profile is determined with $t_{min} = t_{start} + t_{coast} + t_{end}$, where the ramp time periods are determined with $2S_i/\Delta v_i$ and the coast time period with Eq. (2.13).

$$t_{coast} = \frac{\Delta S - \left(S_{start} + S_{end}\right)}{v_t} \tag{2.13}$$

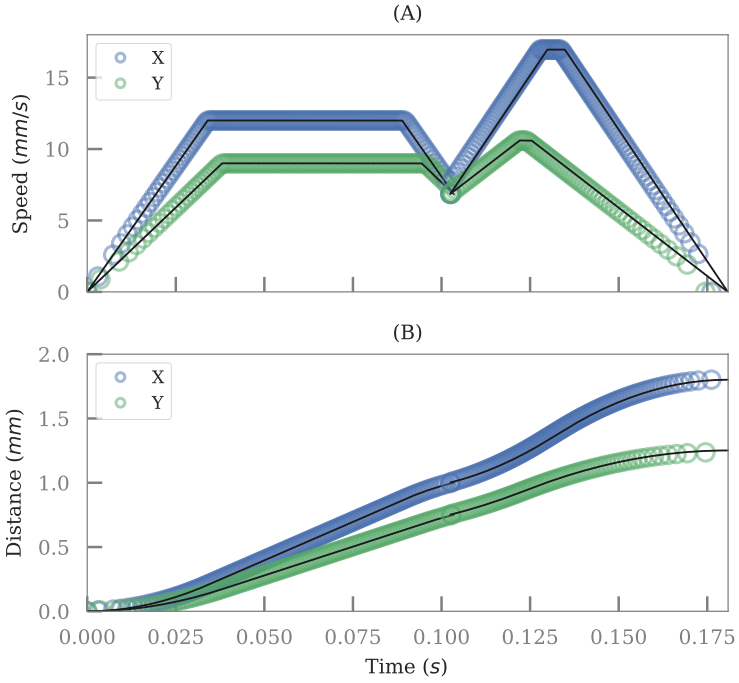


Figure 2.12: Simulated (A) velocity and (B) displacement profiles, for the example G-codes for both the X- and Y-axes. The solid line is the ideal trapezoidal curve and the unfilled circles step pulses.

The primary, or driving, axis is selected as the axis requiring the longest time (t_{min}) to complete the move. This time period is used to recalculate the acceleration, deceleration and the maximum speed for the other, driven, axes. This insures that the move is completed in a co-linear fashion, by slowing down the non-primary axes.

The required acceleration a for a trapezoidal profile is determined with Eq. (2.14). The acceleration end-time is calculated with $t_1 = \Delta v_s / a$ and the deceleration start-time with $t_2 = t_{end} - \Delta v_e / a$.

$$a = \frac{\Delta v_s^2 + \Delta v_e^2}{2\left(v_t t_{end} - \Delta S\right)} \tag{2.14}$$

Note that the acceleration calculations are realised in real time by using a *Bresenham* algorithm in the firmware, as discussed in Section 2.6.4, and that the time points (t_1 and t_2) are either pre-calculated or determined

in real time, for example by checking when the current speed exceeds or equals the target speed for the move.

2.6.3 Real-Time Step Pulse Frequency Realisation

The accurate generation of the required step pulses are critical for precise and smooth printer motion control. A fixed frequency of step pulses creates a constant speed, while constant acceleration or deceleration requires a non-linear change in the stepping frequency or the delay time between step pulses. The step pulse frequency must be determined in real-time and this is realised by the G-code interpreter. The frequency accuracy is limited by the calculation power of the microcontroller and the resolution of the timer, used to generate the set frequency [54, 78].

The current generation of FFF printers are based on 16 MHz, 8 bit ATmega2560 AVR microcontrollers, e.g. an Arduino Mega built into the RAMPS (RepRap Arduino Mega Pololu Shield) printer electronics board. Newer architectures use, for example, an 32 bit ARM processor at 120 MHz (SmoothieBoard). Another option to elevate the processing burden is to use dedicated controllers for each axis. This becomes essential when closed loop control is desired (see [79]). There are already quite a few printers using these faster processors, which can also handle higher order profiles. The 2nd order acceleration calculation is, however, still of relevance.

The delay times (δt) between step pulses are calculated with different methods. First, a pre-compiled data table can be used, but is generally not done, since it requires a lot of memory and cannot be updated in real-time. An example implementation is presented by Jones [57].

The *time-per-step* approximation method, as presented by Austin [78], is a popular and efficient alternative. This method directly determines the approximate delay time per step, by keeping track of the previous time step and an acceleration index, which is the number of steps taken in the acceleration ramp. The previous time step and the acceleration index are combined using an approximation algorithm of the ideal time interval, which is required to achieve the velocity profile.

This *time-per-step* approximation however requires a division. The reciprocal method, *step-per-time*, is claimed to simplify the computational cost, since it does not require any divisions [80, 81]. The *step-per-time* method was realised with a fixed update frequency on an FPGA (Field Programmable Gate Array) by Ranade [80], to control a six axes stepper motor machine with speeds up to 100 kHz (with microstepping). A microcontroller was used to determine the initialisation profile parameters, which were then sent to the FPGA. The algorithm keeps track of the speed time product, given the change in speed at each step (for constant acceleration) and only steps the motor when this product exceeds the actual position.

The RF1000 firmware is based on the *time-per-step* method, but determines the acceleration time interval with $\delta t_C = (atm + vs)/F_{CPU}$, where δt_C is number of CPU ticks before the next step, a the acceleration, t_m the time since the start of the move, v_s the move start speed and

F_{CPU} the frequency of the CPU. Special care needs to be taken with the units and this is presented later. The time interval for constant speed is determined before the move starts.

The calculated time delay, δt, is used to determine the number of CPU ticks before a timer interrupt must be triggered, which calls an Interrupt Service Routine (ISR). The ISR pulses the specific stepper motor drivers. Increasing this time delay will decelerate the motor and decreasing it will accelerate the motor. The number of steps remaining, i.e. the renaming distance, is however independent of the speed. The ISR must therefore determine when the next step pulse will be and determine which axis must be stepped. In this section the *time-per-step* method of [78] is presented first and then the RF1000 firmware approach.

Time Per Step Approximation Approach

The time delay can be calculated, as explained by [54, 57, 78], with the following equations. The delay time between steps δt in seconds is defined with Eq. (2.15) for a speed of f_t in Hertz or for an integer counter c.

$$\delta t = ct_t = \frac{c}{f_t} \tag{2.15}$$

The rotational speed ω is determined with Eq. (2.16), where the motor step angle is $\alpha = 2\pi/n_{spr}$ and n_{spr} is the number of steps per revolution.

$$\omega = \frac{\alpha}{\delta t} \tag{2.16}$$

The rotor speed for any time point t is given by Eq. (2.17a) and the position θ by Eq. (2.17b), where n is the step number, since $\theta = n\alpha$.

$$\omega\left(t\right) = \int_{\tau=0}^{t} \dot{\omega} d\tau = \dot{\omega} t \tag{2.17a}$$

$$\theta\left(t\right) = \int_{\tau=0}^{t} \omega\left(t\right) d\tau = \frac{1}{2}\dot{\omega} t^2 = n\alpha \tag{2.17b}$$

The time delay between step number $n+1$ and n can be defined with Eq. (2.18).

$$c_n t_n = t_{n+1} - t_n = \sqrt{\frac{2\alpha}{\dot{\omega}}}\left(\sqrt{n+1} - \sqrt{n}\right) \tag{2.18}$$

Finally, the expressions for the first (c_0) and nth counter (c_n) values are derived with Eq. (2.19).

$$c_0 = \frac{1}{t_t}\sqrt{\frac{2\alpha}{\dot{\omega}}} \tag{2.19a}$$

$$cn = c_0\left(\sqrt{n+1} - \sqrt{n}\right) \tag{2.19b}$$

The delay counter value determination at time point n requires the calculation of two square roots, which is a costly operation for an 8 bit processor. Note that this calculation needs to be performed in the ISR, while the motion joining calculation, which also has two square roots, is only calculated once per joining point. A Taylor series expansion is therefore used to approximate inter-step delay with Eq. (2.20).

$$cn = c_{n-1} - \frac{2c_{n-1}}{4n+1} \tag{2.20}$$

The approximation introduces an error of 0.44 for $n = 1$. This can be compensated for by multiplying c_0 with 0.676 or by using $c_1 = 0.4056c_0$ [54, 78]. In the simulation presented here the c_1 constant was modified.

A change in the counter delay value results in the acceleration or deceleration of the motor shaft. The acceleration magnitude depends on the constant c_0 and integer number of steps in velocity ramp, n. A change of acceleration, therefore, requires a recalculation of these values.

The speed ramp against time for the X-axis for the first example G-code line is shown in Fig. 2.13. The solid line is the ideal speed ($v_{ideal} = atx + vs$). The rectangles show the step pulse intervals, while the small circles show the ideal step pulse time points and the large circles the approximate points.

The speed at each pulse time point is determined with $v_{exact} = Scnv/\delta t$, where $Scnv$ is the step conversion factor (μsteps/mm). The approximated speed v_{approx} is determined with Eq. (2.20) and unfilled diamonds denote the relative velocity error $ep = \left(v_{approx} - v_{exact}\right)/v_{exact}$. The error peaks at $n = 1$ (7.5 ms) and then rapidly decreases (see [78], for more details on the approximation error).

The rectangles visualise the increasing step frequency needed to increase the velocity with a constant rate. The minimum delay time (δt), which a processor can achieve is therefore the fundamental speed limit, and is determined by the processor speed and algorithm complexity.

The number of steps already taken can be used to determine when to start decelerating. This is computationally very efficient since the firmware only needs to keep track of the current number acceleration steps n and compare this to the number of steps left, D (i.e. the remaining distance). If n is greater or equal to D, then the motor must be decelerated. Note that Eq. (2.20) will result in deceleration if n is negative. Deceleration can therefore be achieved by replacing n in Eq. (2.20) with $i - m$, where $i < m$ and m is the number of steps in which to ramp the speed down to

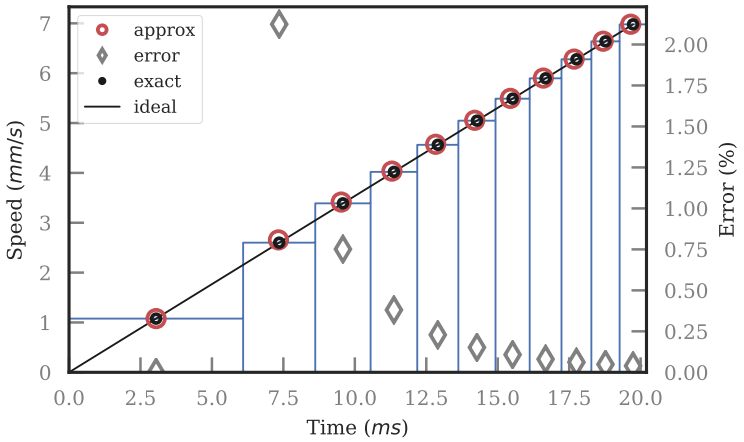


Figure 2.13: Stepper motor pulse timing acceleration with ideal velocity (solid line), exact timer calculation (small circles) and timer approximation (larger, red circles), as well as the approximation error (unfilled diamonds).

zero. The step counter (i) is started at the beginning of the deceleration ramp.

This can be adjusted for non-zero end speeds by predetermining the number of steps required to reach the end-speed, given that the deceleration and maximum speed are known with Eq. (2.12) and by converting the distance in millimetres to number of microsteps, with $1/S_{cnv}$.

Velocity Per Step Approximation Approach

The RF1000 firmware uses a different approach, which is possibly based on the *step-per-time* method presented by Ranade [80], since the current speed is tracked, but with a time interval calculation at each step. The following description is derived from the firmware itself (version 1.10). The source code is not thoroughly documented and this work contributes to the field by explaining the algorithm.

The first calculation step is given in Eq. (2.21).

$$\Delta S_i = \Delta D_i S_{cnv:i} \tag{2.21}$$

This is done after parsing the G-code command to determine the change in distance per axis in millimetres (ΔD_i), as well as the required steps $\Delta D_i = D_{target:i} - D_{current:i}$, where D_{target} is the target and $D_{current}$ is the current axis position, with i representing the X, Y, Z or E axis. The axis with the longest travel distance is selected as the

primary (p), or driving, axis (ΔD_p and ΔS_p). All the other axes must complete their displacement at the time point when the primary axis target is reached.

Next, the fastest time $t_{MinTicks}$ (CPU ticks) in which the primary move can be completed is determined, without using any acceleration, with Eq. (2.22), where F_{CPU} is the CPU clock frequency (tick/s) and V_F the G-code specified feed rate in millimetres per second. The minimum step interval, I_p, is determined with Eq. (2.22b), in ticks per steps.

$$t_{MinTicks} = \frac{F_{CPU}\Delta D_p}{V_F} \tag{2.22a}$$

$$I_p = \frac{t_{MinTicks}}{\Delta D_p} \tag{2.22b}$$

The step interval for each axis, I_i, is determined with Eq. (2.23) and the slowest interval is selected as the primary interval, I_p, where $V_{max:i}$ is the predefined maximum speed for axis i.

$$I_i = \left(\frac{F_{CPU}}{Scnv V_{max:i}}\right)\left|\frac{\Delta S_i}{\Delta S_p}\right| = \left(\frac{F_{CPU}}{V_{max:i}}\right)\left|\frac{\Delta X_i}{\Delta S_p}\right| \tag{2.23a}$$

$$Ip = \max\Big(I_p, I_{X,Y,Z,E}\Big) \tag{2.23b}$$

The primary step interval is used to update the total move time in ticks ($t_{MinTicks}$), which is converted to seconds t_{MinSec} with Eq. (2.24b). This is used to update the interval time and speed for each axis v_i (mm/s).

$$t_{MinTicks} = I_p \Delta S_p \tag{2.24a}$$

$$t_{MinSec} = \frac{t_{MinTicks}}{F_{CPU}} \tag{2.24b}$$

$$I_i = \frac{t_{MinTicks}}{\Delta S_i} \tag{2.24c}$$

$$v_i = s_i \frac{\Delta S_i}{t_{MinSec}} \tag{2.24d}$$

This scales the speed of each axis, but does not consider speed ramping with a constant acceleration. The per axis interval and maximum acceleration $a_{s:i}$ (step/s^2) are used to determine the slowest acceleration, required to reach the target step interval, with $a_{MinTicks} = \min\big(I_i a_{s:i}\big)$.

The acceleration for the primary axis in step/s^2 is scaled with Eq. (2.25a), which limits it so that none of the individual acceleration limits are exceeded. Finally, the step acceleration is converted to a_{sp} and multiplied with a constant ($C = 2^{18}$) in Eq. (2.25b), to enable fast integer calculations in the ISR using bit shifting.

$$a_{sp} = \frac{^{a}MinTicks}{I_p} \tag{2.25a}$$

$$fa_p = \frac{Ca_{sp}}{F_{CPU}} \tag{2.25b}$$

The result of the motion joining optimisation is that the start and end speeds (v_s and v_e) are not necessarily zero. The motion joining routine iteratively adjusts the maximum speed v_{max} of the moves in the print queue. The updated maximum speed is used to recalculate the acceleration and decelerations steps (ΔS_{accel} and ΔS_{decel}) with Eq. (2.26).

$$\Delta S_{ai} = \frac{v_{max}^2 - v_i^2}{2a_{sp}} \tag{2.26}$$

These steps are checked and limited, if the maximum speed cannot be reached, i.e. a triangular profile is required instead of a trapezoidal trajectory.

The ISR uses these predetermined values to determine the next step delay time, by comparing the number of steps taken from the start of the move with the number of steps needed for acceleration and deceleration. If the movement is in the acceleration phase then Eq. (2.27) is used to determine the time delay for the next step pulse δt in steps per CPU tick.

$$v_c = t_m fa_p + v_s \tag{2.27a}$$

$$\delta t = \frac{v_c/2^k}{F_{CPU}} \tag{2.27b}$$

$$t_m = t_m + \delta t \tag{2.27c}$$

The current speed v_c, which is the maximum speed the movement has reached so far, is updated with Eq. (2.27a), where t_m is the time in ticks, since the start of the move.

The current speed is reduced by the number of in-between steps (k) with Eq. (2.27b), which is used to increase the maximum step rate (see

Section 2.6.4). This is done with an efficient bit shift operation, since k is either 0, 1 or 2.

A costly division by the F_{CPU} is required in Eq. (2.27b) to convert the speed to number CPU ticks. The firmware uses an assembler coded lookup-table to approximate this result, which can reduce the accuracy of δt. Finally, the current move time (t_m) is incremented with δt in Eq. (2.27c).

Deceleration is performed if the number of steps left is equal to the number of steps required to complete the deceleration to v_e, as determined by Eq. (2.26). The deceleration calculation is similar to Eq. (2.27), while the pre-calculated time interval is used, if the motion is in coast section (constant speed) of the trapezoidal speed profile.

2.6.4 Trajectory Smoothing

3D printers with stepper motor driven axes need to convert a multi-axis G-code movement command into velocity profiles for each axis ensuring that the movements are completed at the same time. This is done in the Repetier and Gbrl firmware with the Bresenham algorithm, which was developed to control a 2D digital plotter. The algorithm determines which integer points in an XY grid should be used to best represent a line on the grid. A key advantage of this method is that it only requires integer addition, subtraction and bit shifting, making it computationally fast [82]. The algorithm was also later extensively used in computer graphics [83], but it is also applicable to FFF printers.

This method only requires integer counters making it fast and ideal for printer microcontrollers. One issue, however, is that the driven axis can suffer from noise or shaking, called jitter, especially at step frequencies less than 5 kHz. The Grbl firmware implements what it calls Adaptive Multi-Axis Step Smoothing (AMASS) to address this issue. It artificially increases the Bresenham resolution depending on the stepping speed (step frequency). The amount of smoothing depends on the resolution increase, which is set to be more with slower speeds and less with faster speeds. The additional AMASS virtual steps allow the non-primary axis to be stepped, before the driving axis step interval. Aliasing effects are therefore reduced, without sacrificing the accuracy of the Bresenham method, by stepping the driven axis a little earlier [70]. The Repetier uses a similar method and calls it *trajectory smoothing*, by implementing a half-sub step resolution.

The Bresenham algorithm was simulated with a Python script for the first few steps of the first example G-code command in Fig. 2.14. The solid straight line indicates the ideal path, but since the stepper motors can only achieve a finite step size this line must be approximated. The target end position is, however, always reached without any error. The driving axis is plotted on the X-axis for a total displacement of 70 µm. The driven axis has a displacement of 49 µm during this period, with a rounded microstep length of 7 µm.

The step timing for the driven axis is shown in Fig. 2.14(B). The smaller unfilled circles indicate the axis displacement versus the step

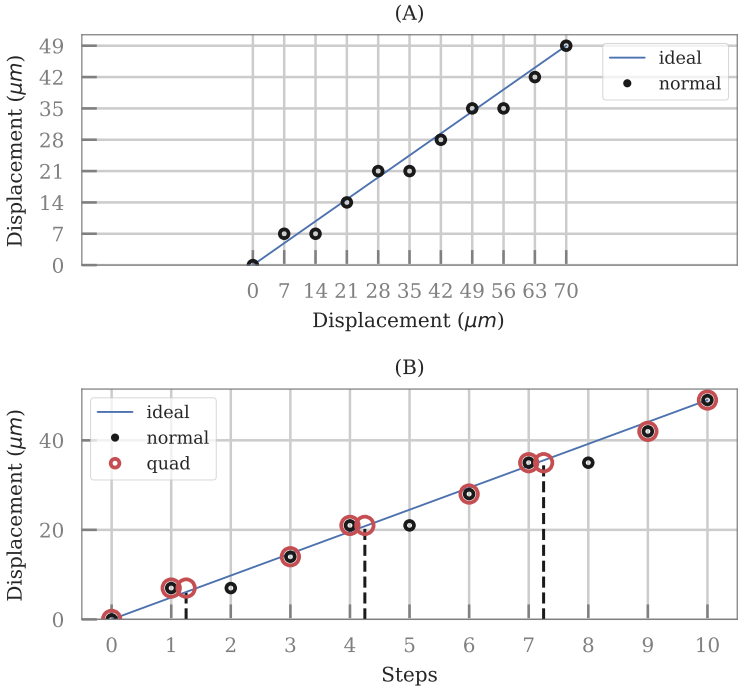


Figure 2.14: Bresenham line approximation algorithm and trajectory smoothing, with (A) the XY step plan in microsteps and (B) the displacement of the driven axis versus time (step number), where the smaller circles are the conventional step locations and the larger, circles the smoothed steps. Vertical dashed lines indicate points were a driven axis were stepped earlier.

number with the conventional Bresenham algorithm. Note that the step number represents time, at a certain step frequency and solid line a constant speed in micrometres per step.

The location of the normal Bresenham points at steps 2, 5 and 8 is brought closer to the ideal line, by virtually increasing the step resolution. In this case a quad sub-step resolution was used, and shows how this method can reduce the risk of jitter at low speeds.

Additionally, the Repetier firmware also increases the number of steps per timer interrupt, if the current stepping speed is greater than a set point frequency. Both double and quad stepping are possible, where quad-stepping is said to increase the maximum step rate to 40 kHz on an AVR based printer board. Note that quad-stepping will decrease the actual

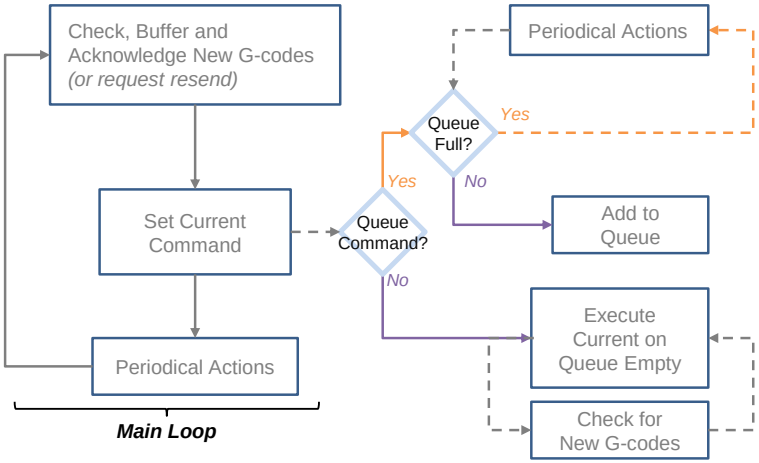


Figure 2.15: Code realisation of command buffer and queue. The main loop is on the left and queue and buffer moves are differentiated on the right-hand side.

number of microsteps with a factor of four [72].

2.6.5 Code Realisation

Flow control is important to reduce any unwanted delays between moves. This is essential for firmware where arcs are divided into small incremental linear extrusions. The host will pause at a specific command until the printer returns a *command acknowledge* response, while the firmware implements a ring buffer to store all the acknowledged commands.

The buffering process for RF1000 firmware is shown in Fig. 2.15, based on [34, 38, 60] and derived from the firmware itself. It uses both a queue and a buffer for G-code commands.

The main loop runs indefinitely, after the printer initialisation (on the left side of the figure), and checks for new commands either from the host-PC or from the SD-card. The printer requests a command resend, if the buffer is full or if a transmission error occurred. A *command acknowledge* response is sent back to the host-PC, if the command is accepted. This establishes flow control between the print server and the printer [60].

Periodical actions are performed after checking for a new command, which includes functions such as reading temperatures and controlling heaters, with a PID control loop.

Additionally, certain commands are also queued to reduce the risk of movement pause or jitter, for example the **G1** command, is placed in the queue. The queued commands have a higher priority than the

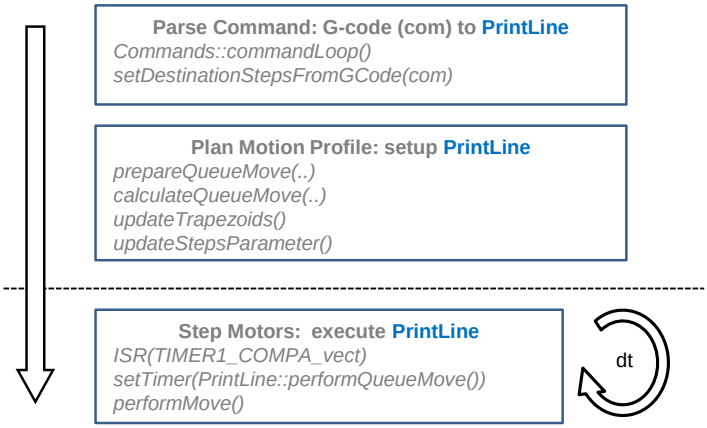


Figure 2.16: Queue move code realisation, where the movement object (*PrintLine*) is first created by parsing the G-code and calculating the trapezoidal trajectory. This also updates the other moves already in the queue. The trajectory is realised in Interrupt Service Routine (ISR), which is triggered after a variable time period (δt). Function names are those used in the RF1000 firmware version 1.10.

buffered ones, but still allow for the periodical actions to be performed. A non-queued command is only executed when all the (currently) queued commands are completed.

Queued motion is realised with a block diagram in Fig. 2.16. A move object (*PrintLine*) for each movement command is created and stored in the queue. The G-code command is parsed in *Commands::commandLoop()* with *setDestinationStepsFromGCode*, which determines the move parameters. For example, a new X-axis target position in millimetres is converted to steps. These update the printer object, which is used by *prepareQueueMove* to determine the change in position required as well as the primary axis, for the Bresenham algorithm (Section 2.6.4).

The print move is further initialised with *calculateQueueMove*, which calculates the target speed for the primary axis as well as the speeds for each axis. The target speed and distance are used to determine the primary axis acceleration and the trapezoidal velocity profile as determined with *updateTrapezoids*. This performs the backward and forward queue calculations needed to determine the maximum move joining speed (Section 2.6.1). This function also calls the *updateStepsParameter* method, which determines the number of steps required for acceleration and deceleration (for the primary axis). A queue move is finally marked as locked if both the start and end speed are fixed.

The queue moves are processed on a first-in, first-out (FIFO) basis, by the Interrupt Service Routine (ISR), which is called every time the step counter delay time is passed. The queue move is processed by the ISR with *performQueueMove*, which calls the *performMove* function. This function finally steps the motors, with the Bresenham algorithm and determines the next interrupt counter period, with Eq. (2.27), for example.

Note that the ISR and motion planning methods are responsible for many other processes as well, such as checking the limit switches, handling a printing pause and performing bed level compensation. The process flow diagram in Fig. 2.16 only describes the main functions called to achieve printer motion and material extrusion.

2.7 Materials and Feedstock

This section gives an overview of selected materials used by Material Extrusion printers. The processing temperatures, cost, printing advantages, disadvantages as well as environmental impact are considered first. The rest of this section focusses on poly(lactic acid), also known as polylactide (PLA), since PLA is used exclusively in this work.

2.7.1 Material Options

Different feedstock materials or filaments are used in FFF for different purposes. These materials require different processing parameters and considerations. The first RepRap printers, for example, used polycaprolactone (PCL), since it has a very low melting point of 60 °C. This was, however, abandoned as the feed mechanism could not get sufficient grip on PCL. Furthermore, PCL also exhibited stringing issues after extrusion. It was therefore replaced by ABS, which has a higher Young's modulus than PCL, is cost effective and commonly available. This change was made possible by improving the feed mechanism and liquefier so that hotter temperatures can be achieved. The ABS filament was then replaced by PLA, for ecological reasons, but it was also found that PLA is easier to print with [30].

Currently, there are a lot of different materials, including composition variations and filament with nano-particles. Particle filled filaments can be used, for example, to realise objects with electronic functionality, as shown by Hampel, Monshausen, and Schilling [84], or the production of metal parts [85].

Other special material functions include: biodegradability, *glow-in-the-dark*, magnetic, colour changing, wax (for investment casting), flexibility and translucency. Special filaments are also available for extruder cleaning and support structure generation [86].

Only a few of the popular or general FFF material types are presented in Table 2.6, where the acronym, short description and application area or examples are given for each material, based on [16, 86–88]. Custom-made materials for industrial printers are not discussed, see for example [28] for description of materials made by Stratasys for their FDM® machines.

Table 2.6: Materials used in Fused Filament Fabrication

Name	Description	Application
ABS: Acrylonitrile Butadiene Styrene	Second most popular after PLA. Tends to shrink: warp and reduces accuracy. Produces less stringing and can be acetone smoothed. Printing requires ventilation.	Tough and durable parts, e.g. tool handles or electrical enclosures.
ASA: Acrylonitrile Styrene Acrylate	Common alternative to ABS, with high UV, temperature, impact and chemical resistance. Warps less than ABS. Needs to cool slowly to reduce cracking.	Outdoor parts which should not degrade (yellow).
HIPS: High Impact Polystyrene	Lightweight material, stronger than both PLA and ABS. Warps less than ABS.	Mainly used as support material for ABS, but can be used in place of PLA.
PA (Nylon): Polyammide	Semi-flexible material, with high impact and abrasion resistance, good strength, flexibility and durability properties. Hygroscopic: requires special storage.	Durable parts: tools or functional prototypes.
PC: Polycarbonate	One of the strongest 3D printing materials, with high heat and impact resistance. Hygroscopic: requires special storage.	Application in harsh environments, can also be used for lighting projects, due to optical clarity.

Continued on next page

Table 2.6 – continued from previous page

Name	Description	Application
PC-ABS	Combined strength and heat resistance of PC, with the flexibility of ABS. Printing is more complex due to higher temperatures, warping and hygroscopic nature. Requires bake-out before use.	One of the most widely used industrial thermoplastics.
PEEK: Polyether ether ketone	High-performance, semi crystalline thermoplastic, with high strength and stiffness as well as high ductility	Sterilization is possible for medical and food contact applications.
PETG: Polyethylene terephthalate glycol-modified	A good middle ground between ABS and PLA: more flexible and durable than PLA, but easier to print than ABS. Hygroscopic: requires special storage.	Parts requiring smooth surface, water and stress resistance, e.g. mechanical or protective parts.
PMMA: Polymethyl Methacrylate (Acrylic)	Hard, scratch resistant, lightweight, with good shatter resistance and transparency. Can be acetone smoothed. Difficult to print, requires a heated chamber.	Industrial applications include headlight lenses and ice-rink glass. Avoid flexure parts, due to inflexibility.
PLA: Polylactic Acid	Most popular filament due to ease-of-use, accuracy and low cost. Biodegradable and environmentally friendly.	Avoid flexure parts, outdoor applications and high temperatures. Applications include prototypes and containers.

Continued on next page

Table 2.6 – continued from previous page

Name	Description	Application
PP: Polypropylene	Tough, flexible, light, chemically resistant and food safe. Difficult to print, due to heavy warping and poor layer adhesion. Possibility of recycling household objects, since many are made from PP.	High-cycle, low strength applications, due to fatigue resistance.
POM: Polyoxymethylene	Used e.g in camera focusing mechanisms. First layer adhesion problematic. 3D prints nearly as functional as mass-produced parts.	Applications requiring a low coefficient of friction.
Filled: Carbon Fibre (CF)	CF reinforced, with PLA, ABS, PETG, or Nylon as base material. Needs a wear resistant nozzle.	High structural strength and low density parts, e.g.: structural applications.
Filled: Metal	PLA or ABS mixed with a fine metal powder (e.g. bronze, copper, aluminium, or stainless steel). Needs a special wear resistant nozzle.	Achieves a unique metallic finish and increased part density. Mainly for aesthetic purposes, e.g. figurines and models. Avoid high stress functional applications.
Filled: Wood	PLA infused with e.g. cork or wood dust.	Only for aesthetic purposes, e.g. architectural scale models.
Flexible (TPE): Thermoplastic Elastomers	Various degrees of flexibility, but hard to print.	Vibration damping or parts with rubber like properties e.g. robot wheels

2.7.2 Material Comparisons

The extrusion processing temperatures for different materials are shown in Fig. 2.17, based on [88] (these temperatures may vary based on specific material compositions). Materials requiring higher processing tempera-

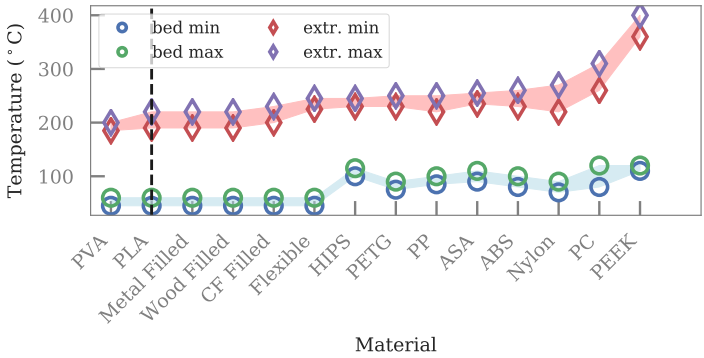


Figure 2.17: Material processing temperatures for FFF materials, with the minimum and maximum bed temperature (unfilled circles) and extrusion temperatures (abbreviated as extr., unfilled diamonds).

tures have, in general, better mechanical and durability characteristics. Hotter processing temperatures, however, require more expensive extruder designs. Another important consideration is the material behaviour after deposition. Industrial Molten Extrusion printers, for example, use both a heated bed and chamber to reduce ABS warping. The bed temperatures for different materials are also shown in Fig. 2.17.

The figure shows that PLA is an interesting material for FFF printers, if processing temperatures are the only consideration. It has both a low bed temperature and a low extrusion temperature, which makes it easier to print. Note that it is also possible to print PLA without a heated bed at all.

The cost of the selected materials are shown in Fig. 2.18, with a scale of $1 - 10$, where 1 is the cheapest and 10 the most expensive, which was calculated by dividing the material cost (taken from [88]) with the maximum cost and multiplying it with ten. PLA, along with ABS, is found to be the second cheapest filament.

The *printability*, as estimated by [88], finds PLA, along with PETG, the easiest to print, this means that considerations such as warping, cracking or special storage are not a major issue, as compared with the other materials.

The mechanical and functional properties of the materials are compared in Fig. 2.19. Note that all these values are subject to material modifications and the figure only provides a qualitative comparison. All the values are based on [88].

The material durability can either be a desired property or unwanted, depending on application. For example, PLA is biodegradable and therefore not suited for outdoor use, but this can also be seen as an advantage.

The maximum operation temperature strongly depends on the mate-

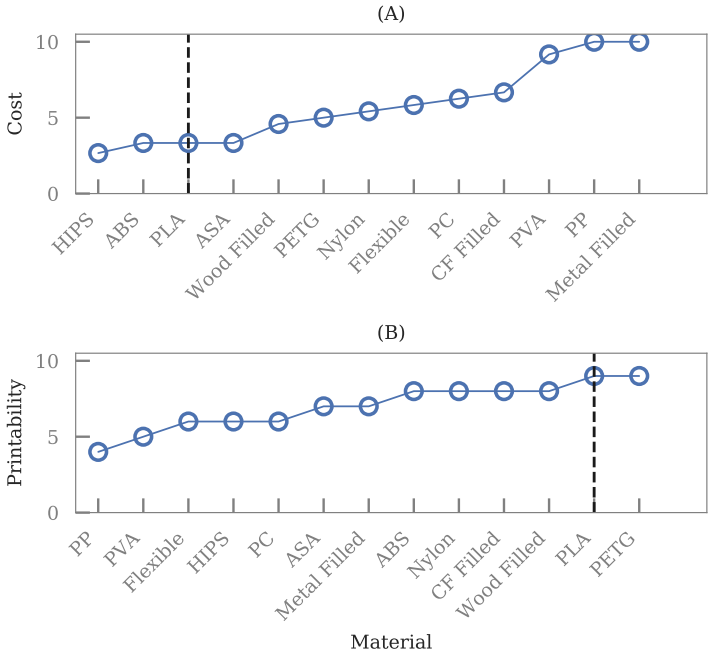


Figure 2.18: Filament (A) cost and (B) *printability* comparison based on [88], where ten indicates the most expensive (A) or the easiest to print filament (B).

rial glass transition temperature. The scale conversion was achieved by dividing each material operating temperature by the maximum and multiplying it with ten. PLA is easier to process with a lower glass transition temperature, but this also means it cannot be used to print a coffee mug, for example.

The strength scaling was determined by dividing each material with maximum ultimate strength and multiplying it by ten, while the flexibility is a relative number given by [88]. The *flexible* filaments are made with a composition of a thermoplastic and elastomer are, as expected, the most flexible. Examples of flexible materials are thermoplastic elastomers (TPE), thermoplastic polyurethane (TPU) and thermoplastic copolyester (TPC). These filaments are, however, difficult to process in general. The relative brittle nature of PLA means that it scores relatively low, compared for example to ABS.

The raw feedstock material is normally granular. This is processed to produce the FFF filament. This additional processing step increases the cost of the material. Liquefier designs which can process the granular

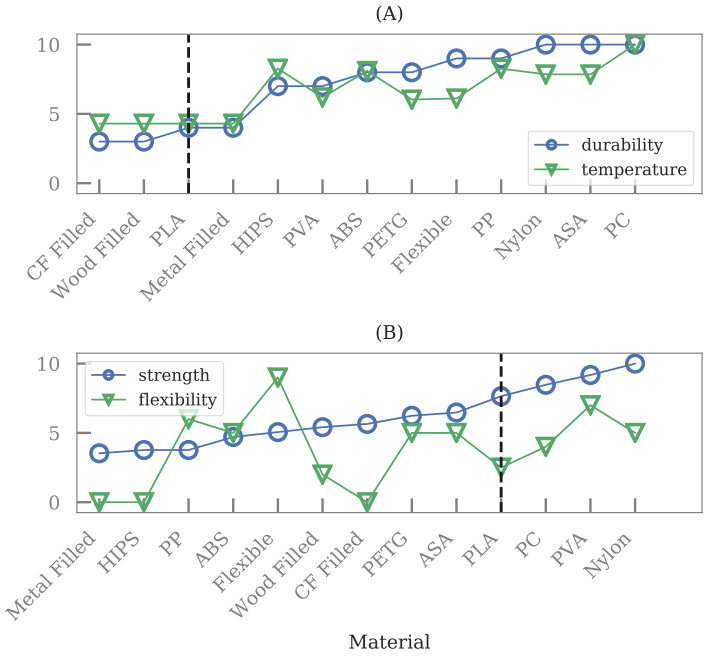


Figure 2.19: Material (A) durability (unfilled circles), operation temperature (unfilled triangles), (B) strength (unfilled circles) and flexibility (unfilled triangles) comparison, with ten the best and zero the worst.

feedstock are therefore considered, by for example [85, 89]. This, however, does not seem to be a general trend. One advantage of using filament is the possibility of printing in space, where the use of powders or liquids are more complicated (see [16]).

2.7.3 Environmental Impact

The FFF process produces ultra fine particle emissions, where the amount of particles and the size of the particles depend on the material and the temperature, according to various emission studies [90–92]. It is recommended that a printer enclosure, with air filtration, is used for ABS, while it is not essential for PLA. A well ventilated room is, however, recommended as well as reducing the processing temperatures [92].

The environmental impact and sustainability are also material considerations, independent of how 3D printing diffuses into the market. It is important to consider the material impact on the environment and health of humans. The initial work by Jones *et al.* [30] already cites PLA as a

more environmental friendly solution and suggest that 3D printing with PLA can even have an environmentally positive impact.

Polymer pollution is a serious global issue, since cheap petroleum derived synthetic polymers require a long time to fully degrade. Such polymers also depend on non-renewable crude oil reserves and this is motivating research and development of polymers derived from renewable sources. A fully biodegradable polymer, with a carbon credit, is crucial for a sustainable future [93].

2.7.4 Polylactic Acid

PLA, a thermoplastic polymer, is one of the first commodity polymers produced from annually renewable resources, such as corn or sugarcane. The environmental benefits of PLA include reduced green house gas yield and low energy requirements for production. Sustainable agriculture are, however, also important, even if the feedstock material is annually renewable, to ensure food security and to reduce deforestation, green house gas emission (through the use of fertilizers) as well as soil erosion [93].

PLA was first synthesized in the 1800s and production was investigated as early as 1932. Initially the production was inhibited by high manufacturing costs. That was until a new company was formed in 1997, which focussed on the production of PLA and made it a large-volume, popular and biodegradable plastic [93, 94].

The material properties determine the processing requirements and limits. Important properties are the extrusion temperature and melt flow characteristics. The processing temperatures, along with the heat capacity and conduction of the material, determine the liquefier specifications, for a specific volumetric flow rate.

A PLA grade intended for the production with FFF is *Ingeo Biopolymer 3D850* [95], by *NatureWorks*, which is one of the leading PLA producers [93]. It has a melting temperature (Tm) range of 165 – 180 °C and glass transition temperature (Tg) of around 55 – 66 °C. The material will transform from a glassy to a rubbery state if it is heated above Tg, and finally to into a viscous fluid around Tm [94].

The FFF feed mechanism must also generate enough pressure to force the molten filament through the nozzle. The required force depends, for example, on the geometry, flow rate, material friction coefficient and melt viscosity, η [96].

The viscosity of water is, for example, 10^{-3} Pa · s, while the viscosity for extrusion type polymers are around 10^{2} – 10^{5} Pa · s [97]. Polymer viscosity is one of the most important material processing parameters, since it determines the resistance to flow [98, 99]. Rheology, the study of the deformation and flow of a fluid, is used to determine viscosity and the dependence of it on temperature as well as shear rate. These relationships are critical to understand the PLA processing requirements.

Material suppliers, however, usually only supply the MFR or MFI (Melt Flow Rate or Melt Flow Index) as an indication of the material viscosity.

The measurement procedures to determine the MFR are standardised and involve the measurement of mass flow per time unit through an orifice, as a result of an applied load. A high MFR indicates a low molecular weight and a low viscosity polymer[97]. The MFR for the Ingeo Biopolymer 3D850 is 7 – 9 g/10min [95], while MFR for NatureWorks PLA *Grade MAT2238* is 8 g/10min [100].

Furthermore, the PLA melt is, like many other thermoplastics, a pseudoplastic, non-Newtonian fluid [97–99]. Detailed rheological characterisation for PLA is required in order to predict the extrusion requirements, but this information is in general not supplied by the filament producers. Nor do they state which granular feedstock material was used to extrude the filament. They usually only supply T_m and the filament diameter.

NatureWorks PLA *Grade MAT2238* [100] is often used in FFF literature, since the freely available test report [100] provides detailed property values. As an example, the first fully numerically resolved simulation of FFF used this report [101]. The datasheet [95] and the measurement report [100] are also used in this work for baseline PLA filament values.

Viscosity Models

A list of viscosity models applied in FFF literature are given in Table 2.7 and discussed next.

Table 2.7: Non-Newtonian inelastic viscosity models used to model flow in FFF extruders. [102] used PZT/ECG9, a thermoplastic carrier infused with PZT powder.

Model	Material	Reference
Power-law	PCL	[103, 104]
Power-law and Arrhenius	PZT/ECG9,PCL	[102, 105]
Cross-WLF	PLA, ABS	[101, 106]
Carreau	ABS	[65]

A Newtonian fluid exhibits a linear ratio between shear rate and shear stress, with $\eta = \tau/\dot{\gamma}$, where the shear rate $\dot{\gamma}$ is the rate at which fluid layers move past each other and τ is the shear stress.

The simplest *non-Newtonian* fluid model for viscosity is the power-law model given by Eq. (2.28), where the viscosity is the gradient of a log-log plot of the shear strain and shear rate, with K the consistency index and n the flow index. Note that a fluid is Newtonian if $n = 1$. This model can also be used for PLA, if the shear-rate is larger than $10\,\mathrm{s}^{-1}$ [107].

$$\eta\left(\dot{\gamma}\right) = K\dot{\gamma}^{n-1} \tag{2.28}$$

The viscosity decreases with the power-law model with increasing shear rate, if $n < 1$. This is called shear-thinning or a pseudoplastic flow. The flow index indicates the deviation from ideal Newtonian flow. An increase in the consistency index K will result in an higher viscosity, which will require a greater extrusion force. The flow index, for example, for polycarbonate (PC) is approximately 0.8, but for rubber it is around 0.2 [97].

The flow behaviour of PLA is a combination between irreversible viscous flow and reversible elastic deformation. The elastic and viscous flows can be combined to form a complex viscosity model [99].

The viscous flow is due to polymer chain slippage and the elastic deformation due to molecular entanglement. However, as shear rates increase the molecules orient, which reduces the number of entanglements, allowing the polymer chains to flow more easily past one another into narrow cavities. The molecular chains are also broken down by the shearing action during extrusion, i.e. being forced through the nozzle, decreasing the viscosity [98, 108].

The viscosity also decreases at hotter temperatures, due to the higher kinetic energy of the molecules, which weakens the connections between the molecular chains [108].

The temperature dependence of the PLA melt viscosity can be modelled with an Arrhenius equation Eq. (2.29), where η_0 is the zero-shear viscosity, R is the gas constant ($8.3144\,\mathrm{J \cdot K/mol}$), T the temperature in Kelvin, A a constant ($\mathrm{Pa \cdot s}$) and E the flow activation energy (J/mol), which is a measure of the energy required to overcome the internal attraction forces and resistance, due to entanglement. A larger E indicates that the viscosity has a higher temperature sensitivity [107–109].

$$\eta_0\left(T\right) = Ae^{\frac{E}{RT}} \tag{2.29}$$

The Cross and Carreau-Yasuda models are frequently used to fit data over the entire shear rate range, improving on the power-law model [97, 99].

The material testing report [100] provides parameters for the *Cross-WLF Viscosity Model*, which combines the Williams-Landel-Ferry (WLF) temperature dependence model and the Cross shear rate dependence model [108].

The Cross-WLF model is given in Eq. (2.30) [108].

$$\eta\left(\dot{\gamma}, T\right) = \frac{\eta_0}{1 + \left(\frac{\eta_0 \dot{\gamma}}{\tau^*}\right)^{(1-n)}} \tag{2.30}$$

The critical shear stress is τ^*, which is at the transition point between the Newtonian and non-Newtonian viscosity, with K the Cross model time constant equal to η_0/τ^*. The logarithmic downward slope is equal to $n - 1$, where n is the power-law flow index [110].

The zero shear viscosity η_0 is determined with Eq. (2.31c), where D_1 is the viscosity at the reference temperature D_2, while A_1 and A_2 describe the viscosity temperature dependency.

$$T^* = D_2 + D_3 P \tag{2.31a}$$

$$A_2 = A_2 + D_3 P \tag{2.31b}$$

$$\eta_0 (T) = D_1 e^{\left[\frac{-A_1 \left(T - T^*\right)}{A_2 + (T - T^*)}\right]} \tag{2.31c}$$

Temperature therefore shifts the graphed viscosity up- or downwards, depending on A_1 and A_2, while η_0 determines the initial plateau, before shear-thinning starts at shear rate K [108, 110].

The model parameters for NatureWorks Grade-MAT2238 PLA [100] are given in Table 2.8.

Table 2.8: Parameters for NatureWorks Grade-MAT2238 PLA Cross-WLF Viscosity Model [100]

Coefficient	**Value**	**Unit**
A_1	20.194	-
A_2	51.600	K
D_1	$3.317\,19 \times 10^9$	$\mathrm{Pa \cdot s}$
D_2	373.15	K
D_3	0	K/Pa
n	0.2500	-
τ^*	$1.008\,61 \times 10^5$	Pa

An FFF extruder is similar to capillary rheometer [111], which is used to measure material viscosities by forcing a molten material through an orifice. The pressure driven flow will have a velocity gradient and the shear rate will be a maximum at the wall and zero in the centre. This makes the flow non-homogeneous.

A capillary rheometer can therefore only be used to measure apparent viscosity $\eta(\dot{\gamma})$, and is frequently used to measure high shear rates ($\dot{\gamma} > 100\,\mathrm{s}^{-1}$), which are important for processes such as molten polymer extrusion [110].

The apparent Newtonian shear rate at the capillary rheometer wall $\dot{\gamma}aw$ is determined with Eq. (2.32a), where $\dot{Q}$ is the flow rate and R the radius. It is corrected for the shear thinning behaviour of the melt (the non-parabolic velocity profile), using the Weissenberg-Rabinowitsch correction with Eq. (2.32b), where τ_w is the wall shear stress [110].

$$\dot{\gamma}aw = \frac{4\dot{Q}}{\pi R^3} \tag{2.32a}$$

$$\dot{\gamma}w = \frac{1}{4}\dot{\gamma}aw\left(3 + \frac{d\left(\ln \dot{Q}\right)}{d\left(\ln \tau_w\right)}\right) \tag{2.32b}$$

The $\dot{\gamma}w$ can be reduced to Eq. (2.33), since $(d\ln\dot{Q})/(d\ln\tau_w) = 1/n$ for power-law fluids [112].

$$\dot{\gamma}w = \dot{\gamma}aw\left(\frac{3n+1}{4n}\right) \tag{2.33}$$

The viscosity of NatureWorks Grade-MAT2238 PLA as described by the Cross-WLF model (Eq. (2.31)) is shown in Fig. 2.20, with parameters according to Table 2.8.

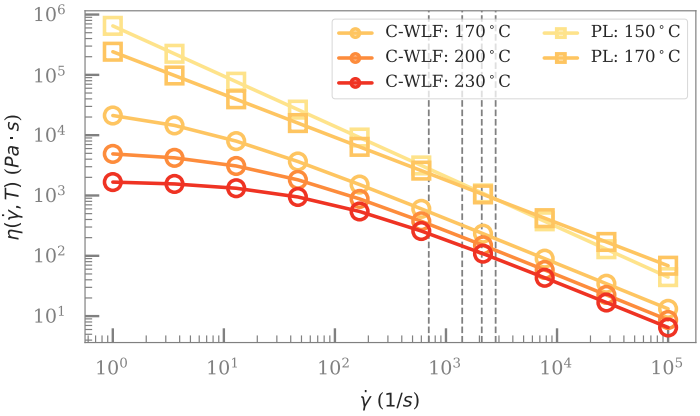


Figure 2.20: Comparison of some of the available models for PLA viscosity dependency on shear rate and temperature, unfilled circles indicate the Cross-WLF (C-WLF) model for NatureWorks Grade-MAT2238 PLA [100] and unfilled squares the power-law (PL) parameters for amorphous PLA [98].

The predicted logarithmic decrease of viscosity with increase of shear rate is clearly seen. The model is graphed for three temperatures, 170 °C, 200 °C and 230 °C, which shows how the curve is shifted upwards with decreasing temperature.

The power-law models of Fang and Hanna [98] are also shown for amorphous PLA at 150 °C and 170 °C for reference.

Normal to fast print speeds for the RF1000 printer (25 mm/s, 50 mm/s, 75 mm/s and 100 mm/s) are converted to shear rate with Eq. (2.33), using n equal to 0.25. These speeds are the vertical dashed lines in the figure. This shows that it is expected that the RF1000 extruder is operating in the shear-thinning range of PLA.

The predicted viscosity, by the Cross-WLF model, at an extrusion temperature of 170 °C and a print speed v_p of 50 mm/s is 3.2×10^2 Pa · s at 1400 s^{-1}, with a nozzle radius R of 0.5 mm/s, $Q = v_p A_n$ and A_n the cross-sectional area of the nozzle.

The power-law parameters predict the viscosity at the orifice wall ($\eta(\dot{\gamma})$) as 1.4×10^3 Pa · s, with the same conditions, which is significantly more. Nevertheless, these models provide useful values for Section 2.8, which discusses the extruder design.

2.8 Extruder Design and Modelling

The main goals of FFF extruder design can be defined as the increase of (1) printing speed (2) resolution (3) number of materials, (4) reliability and (5) control, whilst reducing (6) cost.

Printing speed and resolution are inversely proportional to each other. The simplest way to print things faster is to increase the nozzle diameter, however, this will reduce the resolution. A faster printing speed also requires faster feed rates, and assuming constant viscosity, larger feed forces. Increasing the filament diameter, which acts as the extrusion piston, will allow for faster extrusion rates, but will also increase the time required to melt the filament, increasing the liquefier length and cost.

The print time decrease ratio r_t, as result of an increase in the nozzle diameter D_i, can be estimated with $r_t = dt_0/dtn = (D_0/Dn)^2$, where $dt_i = V_p/\dot{Q}_i$ is the time needed to print a volume V_p and $\dot{Q}_i$ is the constant volumetric flow rate, for nozzle sizes 0 and n. This assumed that both prints run at the same constant speed v_p, which gives $\dot{Q}_i = v_p A_i$, where A_i is the cross section equal to $\pi D_i^2/4$.

A time increase of 25 times is predicted with this simplified model, if a nozzle diameter of 0.5 mm is replaced with a 0.1 mm nozzle.

A detailed FFF rate limit analysis of the three main printer modules, the material extruder, filament heating and the positioning system, is performed by Go *et al.* [65]. The maximum throughput rate of each module determines the fastest build rate or print time. Material extrusion and heating is found to limit the feed rate to approximately 9 mm/s, with a maximum extrusion force of about 60 N.

These values converted to the printer configuration used in this work are equal to a maximum print speed of 0.3 m/s ($vp = v_f(D_f/Dn)^2$) and a maximum liquefier pressure of 9.4 MPa ($P = F/A_f$).

The quickest print time is, however, not the only design goal. The part strength, for example, is another important aspect. This depends on the layer height [113], which is strongly influenced by the nozzle diameter [114, 115]. The part dimensional accuracy is also strongly related to the nozzle diameter [116, 117].

The number of materials can refer to both in-process options as well as to the number of materials which can be successfully processed with the same printer. Increasing the number of materials a printer can process will also increase system complexity, in order to cater for different material processing requirements, which will reduce printer reliability.

Finally, the ability to achieve fine control will increase the attainable resolution and reliability as well as the printing speed, but this also increases the cost. Cost can be defined as production cost, but also as the weight or the accessibility (complexity) *cost*. Remember, that the ideal FFF extruder is lightweight (Section 2.4.4) and accessible to a wide range of users (i.e. the maker community) [30].

The short discussion above aims to show *why* careful extruder design is important, since all of these factors are directly influenced by how well the extruder melts the thermoplastic, feeds the filament and extrudes the melt. Developing an understanding of these processes through modelling and experimentation is therefore required to explain *how* an FFF extruder works.

This section first describes different types of extruders as well as the RF1000 extruder, after which different models are reviewed for the liquefier pressure drop, temperature distribution as well as the feed mechanism slippage. The models are also applied to the printer configuration used in this work to illustrate the model predictions.

2.8.1 Types of Extruders

General extruder categories are shown in Fig. 2.21. The plunger or syringe type is typically not used for thermoplastics, but rather for paste, gel or clay materials. See, for example, the works by Temming [118] or Amza, Zapciu, and Popescu [119], where the design of such an extruder is discussed.

The screw type extruder uses pellets or granular material instead of filament. This is based on the industrial extrusion technique. A screw type extruder for FFF is presented by Valkenaers *et al.* [121] as well as by Giberti, Sbaglia, and Silvestri [85] and Drotman *et al.* [120]. A patent application by Zinniel and Batchelder [122] also discusses a "conical viscosity pump", which is a type of adjustable height screw extruder.

One of the main advantages of this method is that the cost of raw thermoplastic granular material is significantly less than that of the processed filament. Another advantage is that it does not suffer from feed slippage, as compared to the filament extruders. The extrusion screw, however,

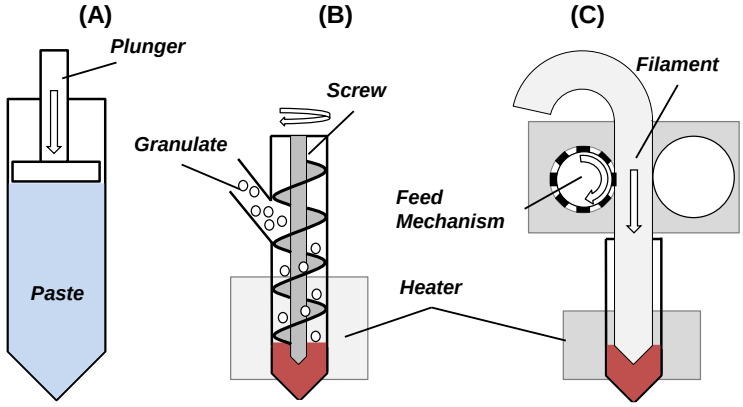


Figure 2.21: Extruder types: (A) paste extruder with a plunger or syringe (B) a screw type and (C) a filament extruder. The screw type diagram is based on [120].

adds significant weight to the print head and must be kept stationary or only moved in the Z-axis (axis of minimum motion). This, however, increases the system cost and complexity.

Furthermore, a filament type feed mechanism can be manufactured by the *maker* community by milling a bolt or using parts from a RC (Remote Controlled) servo motor, for example, while an extruder screw is more complex to produce.

The filament type extruder was introduced in Section 2.1 and it is a more popular method. It uses the filament itself as the piston to force the molten material through the nozzle. Multi-material printing can be realised by using either more than one extruder [39] or with a *mixing nozzle*, where there are more than one feed mechanism supplying a singular nozzle and liquefier.

Full-metal filament extruders are becoming more popular and can print at higher temperatures, with a metal construction and active cold-end cooling. The extrusion principles are, however, similar to the plastic type extruders.

2.8.2 Types of Filament Feeding Mechanisms

The feeding mechanism for the filament type extruder can be a direct, indirect or a conveyor belt drive.

The conveyor belt feed mechanism is less popular. The main idea is to increase the contact surface area between the feed mechanism and the

filament to reduce feed slippage. The first RepRap polymer transport mechanism, for example, used four pairs of pinch wheels to increase grip [30].

A direct drive is shown in Fig. 2.21. The feeding mechanism is moved with the nozzle, allowing for maximum force transfer and pressure generation. A gear or knurled shaft is mounted on the motor shaft. This gear can be in direct contact with the filament or it can drive an additional set of gears, which feed the material. A pinch gear or pulley is also needed so that the driving gear can grip the filament.

The feed mechanism can either be a pinion and rack configuration (as shown in the diagram), where the filament becomes the rack as the pinch bearing forces it into the feed gear, or a worm gear configuration. The patent by Batchelder and Swanson [123] describes such an approach, which increases the contact surface area, which improves the feed mechanism grip.

The filament type extruder can be made light-weight by using an indirect drive or *Bowden* design, by removing the extruder motor from the print head and feeding it through a low-friction tube to the nozzle. This can reduce the space requirements for the print head as well as improve movement dynamics. The cost is, however, a more complex drive train and an increased distance between drive and liquefier, which reduces the extrusion force.

This design can be modified to overcome friction issues (for example, if a long tube is used), by using two motors to feed the filament, with a smaller motor attached to the extruder. This, again, increases the complexity and risk of failure.

Different feed mechanisms and liquefiers can be combined. The patent [124], for example, depicts a combination of a conveyor belt feeder combined with a screw extruder.

This work focusses on single filament feed extruders with direct drives, but most of the concepts can be generalised to other techniques.

2.8.3 The RF1000 Liquefier

The RF1000 uses a classic single filament feed extruder with a direct drive feed mechanism and a pinch pulley, as shown in Fig. 2.22. The photo of the RF1000 hot-end (V2) shows the inlet, cold-break, liquefier (heater and thermistor) as well as the nozzle components. Two of the four wires are for the heater element and the other two for the thermistor. The cold-break is made from a high temperature machinable plastic and the nozzle from brass. The nozzle is screwed onto the hot-end and can be exchanged.

The V2 hot-end is available in both a 3 mm and 1.75 mm filament diameter version. This scales the internal diameters, but it is assumed that lengths and external diameters remain the same.

A representative CAD model of the V2 liquefier and nozzle assembly is shown in Fig. 2.23. The dimensions were derived from measurements of the hot-ends available and mentioned values are only representative,

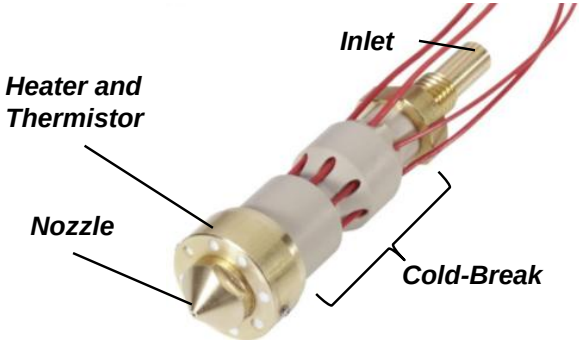


Figure 2.22: RF1000 V2 hot-end.

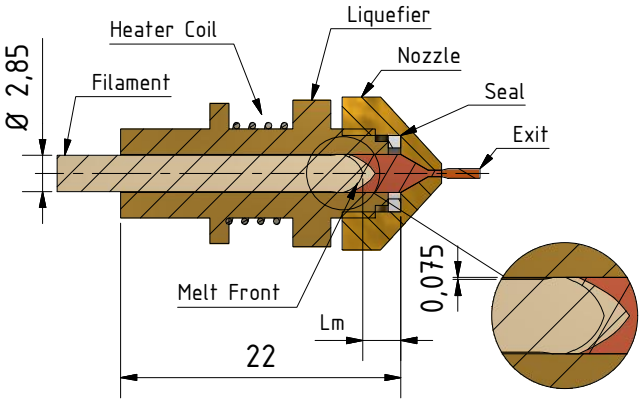


Figure 2.23: CAD of the RF1000 V2 (3 mm) hot-end. The solid filament enters the liquefier on the left, gradually heats up, melts and is pressurised out of the nozzle. A detail bubble conceptualises the melt front profile (dimensions in mm).

i.e they are not a part specification. Realistic values are, however, needed to model the liquefier.

The filament diameter, D_f, is slightly less than then the inner diameter of the liquefier, D_L, creating a gap between them of (0.075 ± 0.050) mm, since the filament diameter used in this work is specified to have a tolerance of (2.85 ± 0.10) mm.

The distance from the feed mechanism filament contact point to the

average melt front location is defined as L_0. It is possible that a certain length of material, defined as Lsm, is not yet fully melted, but already touching the liquefier wall. This can be due to filament straightness deviation, insertion angle error or thermal effects such as material expansion or deformation due to softening (before melting).

The melt length L_m is defined here as the distance from the average melt front location to the start of the nozzle conical section, which has a length of L_{nc} and a cone angle of β_n.

Values for the dimensions are given in Table 2.9, where most of the dimensions were measured with a vernier calliper.

Table 2.9: RF1000 extruder dimensions.

Dimension	**Size** mm	**Dimension**	**Size** mm
L_0	80.0	L_{nc}	3.5
L_L	22.0	L_n	1.0
L_{sm}	2.0	D_m	3.0
L_m	4 – 6	D_n	0.5
D_f	2.85		

The melt length, L_m, was estimated by rapidly reversing the filament out of the liquefier (pulling it out), after allowing it to melt, and measuring clearly melted section. This was done a few times after cutting of a length of filament and reinserting it. The semi-molten length, L_{sm}, is approximated as the range of the measured L_m lengths (maximum minus minimum). Note that the extruded material will also increase in diameter after passing through the nozzle. This is called die-swell [115].

The lengths for the liquefier L_L, nozzle conical section L_{nc} and nozzle tip L_n as well as the diameter of the liquefier D_L and nozzle D_n are also given in Table 2.9.

2.8.4 The RF1000 Feeding Mechanism

The RF1000 feeding mechanism uses a pinch bearing and gear to feed the filament, as shown in Fig. 2.24. The feed gear is fixed to the stepper motor shaft. Four compression springs force the freely rotating pinch bearing against the filament, which forces (pinches) the filament between the feed gear and the bearing.

The pinch force should be sufficient so that the feed gear can grip the filament; conversely it should also not be too strong. This is especially an issue with *soft* filament materials.

The gear tooth profile can be optimised to improved grip, see for example [125], where a profile which maximises grip is presented. The

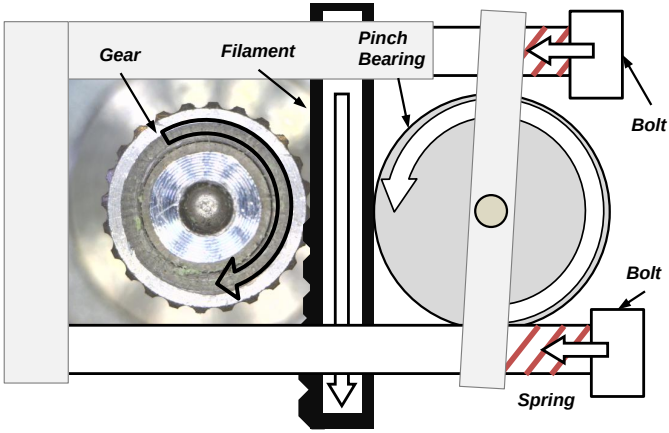


Figure 2.24: Diagram of the RF1000 pinch roller feed mechanism

feed gear should, however, only press into, and not cut, the filament. Digging into the filament will generate shreds, reduce the volume flow rate and potentially break the filament.

The amount of force exerted by the pinch bearing on the filament Fp can be adjusted by using springs of different stiffness (k_{spring}) or by adjusting by rotating the bolt, i.e. increase or decrease the compression length x_{comp}. The pinch force can be determined with $Fp = \left(l_0 - x_{comp}\right) k_{spring}$, where l_0 is the initial spring length [126].

2.8.5 Maximum Feeding Force

The maximum feeding force, which the stepper motor can exert to push the filament into the liquefier, gives an indication of the maximum pressure drop, which the extruder can generate. This in-turn will determine the maximum feeding rate and minimum liquefier temperature, if all the other variables are kept constant. The RF1000 uses the same stepper motor for all axes, including the E-axis, with a holding torque of approximately $48 \times 10^{-3}\,\mathrm{N \cdot m}$ and a gear pitch diameter D_{feed} of 7.276 mm (see Section 2.4).

The maximum force F_{max} can be determined with Eq. (2.34) [115], where τ_f is the motor torque which is equal to 13 N.

$$F_{max} = \frac{2\tau_f}{D_f} \tag{2.34}$$

The pressure which this force generates is 21 MPa, since $P_{max} = F_{max}/A_f$, where are A_f is the filament area. This force, however, should

be reduced with a safety factor of at least two (see Section 2.4), which gives a maximum feed pressure of 10 MPa (assuming no feed slippage). This value agrees well with the maximum pressure of 9.4 MPa, determined previously from maximum force found by the rate limit analysis of Go *et al.* [65].

Filament buckling between the feed mechanism and the melt front also limits the maximum feeding force, especially for softer filaments. Euler buckling (Eq. (2.35)) can be used to approximate the critical pressure (P_{cr}), which can be applied to the filament before it will buckle, where L_{0u} is the length of unsupported filament between the feed mechanism and the melt front [115, 127]. Note that the heat break can support the filament, before it reaches the liquefier.

$$P_{cr} = \frac{\pi E D_f^2}{16 L_{0u}^2} \tag{2.35}$$

Buckling is not seen as problem for PLA, which has an elastic modulus of around 3.5 GPa [100]. The length L_{0u} for the RF1000 extruder is approximately 10 mm, which gives a critical pressure of 175 MPa.

The maximum unsupported length can also be determined by reordering Eq. (2.35). This length is equal to 42 mm, if a pressure of 10 MPa is applied. Note that it is about half of the total distance between the estimated melt front location and the feed mechanism, L_0, which is equal to 80 mm (see Table 2.9). It is possible, therefore, that the filament buckles and presses against the heat break internal geometry during extrusion.

2.8.6 Nozzle Design

The FFF nozzle is normally provided with a thread so that it can be disconnected from the liquefier for a fast diameter change or for maintenance purposes. Important nozzle design parameters are the diameter D_n, tip external diameter, total length, angle of the conical section β and the nozzle material. The total nozzle length includes the tip length L_n, the conical section length Lnc as well as a section, which has the same diameter as the liquefier. The mounting thread can either be external or internal, depending on the liquefier design. The RF100 liquefier uses an external thread.

Four nozzle designs are presented in Figs. 2.25 and 2.26 to illustrate the discussion. All the nozzles are designed for a 3 mm filament diameter. Nozzle number (B) is based on the RF1000 V2 with a 0.5 mm diameter nozzle, while the other designs are based on open source designs, specified by [128]. Nozzle (A) and (B) have a 60° cone angle, with L_n approximately equal to 1 mm. Nozzle (A) has an external thread and is longer than (B), which has an internal thread.

Nozzle (C) and (D) are 0.4 mm diameter nozzles, with cone angles of 100° and 90°. Nozzle (D) also has a two reduction sections. The minimum distance between the external and internal cone is shown (rounded to

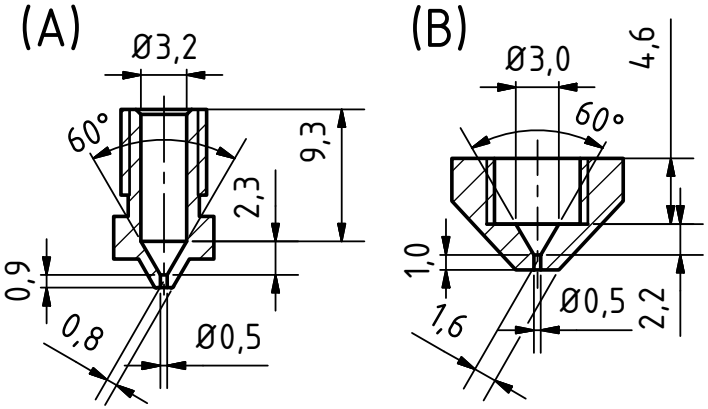


Figure 2.25: Example nozzle designs for a 3 mm filament with a 0.5 mm nozzle diameter: (A) is an external thread nozzle based on [128] and (B) is an internal thread nozzle based on the RF1000 V2 extruder.

0.1 mm) and nozzle (C) is found to be the thinnest, with only 0.5 mm of material.

The nozzle design is determined by the filament material, volumetric flow rate, print resolution and accuracy specifications. The nozzle will also wear during printing, especially if carbon filled materials are used. Requirements for a nozzle can be defined as (1) withstand the melt abrasion and extrusion pressure, (2) conduct heat to the melt, (3) reduce pressure drop, (4) interface with the deposited track and (5) fulfil special requirements.

Wear will be discussed first. Most FFF nozzles are made from brass, since it is easy to machine, low-cost and has relatively high thermal conduction. Stainless steel, hardened steel, aluminium and coated nozzles are, however, also available. Furthermore, tungsten and ruby tip nozzles can be used for very abrasive materials. Printing for medical or food applications can be achieved by using a stainless steel nozzle.

Brass and aluminium have higher thermal conductivity than stainless steel. The nozzle thermal conductivity affects how fast the heat loss due the material flow out of the nozzle can be recovered by the heat flowing from the liquefier. This will in-turn affect the pressure drop, which is also influenced by the material friction coefficient. The heat from the heated liquefier should, therefore, flow as efficiently and effectively as possible to the melt, through the unheated nozzle. This becomes more important at faster flow rates and with longer nozzles.

Additionally, clogging can occur if the deposited track cooling fan

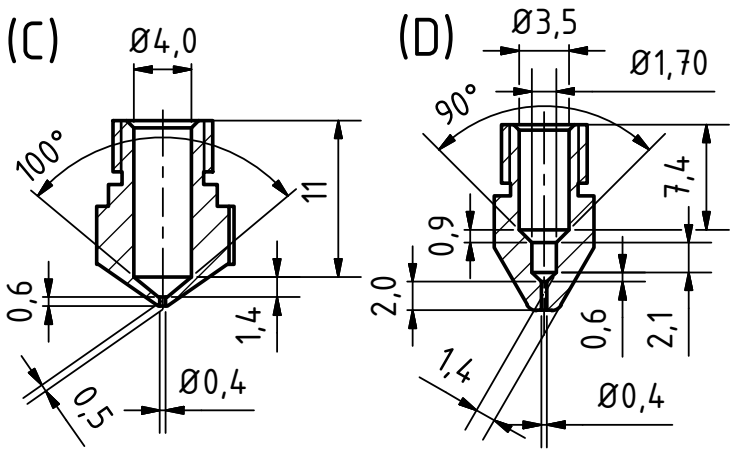


Figure 2.26: Example nozzle designs for 3 mm filament with a 0.4 mm nozzle diameter. Both nozzles have external thread mountings and are based on [128]. Nozzle (D) has two conical reduction sections.

cools the nozzle. Alternatively, build deformation can occur if the nozzle reheats the deposited track. The nozzle should therefore conduct the heat from the liquefier to the melt and not to the surroundings. This can be achieved with silicone rubber hot-end covers, produced and sold by [129]. This isolates the hot-end from the print as well as the cooling fan.

Special coatings or surface finish can be applied to the internal nozzle geometry to improve heat conduction, reduce friction and increase wear resistance (see for example nozzles produced and sold by [128]).

Reducing the heat transfer to the track will improve the stability of tall and thin or overhanging features. This can be done by reducing the tip external diameter. Nozzle (B) has the widest tip at 1.5 mm, nozzle (C) is the thinnest with 0.6 mm and the other two are in between at roughly 1.0 mm (note: these dimensions are not shown in the drawings). A sharper tip is therefore used to create finer features, but will also be less robust and deforms easily if crashed into the bed [130].

On the other hand, a flat (as compared to sharp) nozzle will flatten out blobs and allow for wider tracks. Some deposited material may, however, accumulate on a wider nozzle, which can cause serious build issues later on.

Interestingly, a patent application by Swanson *et al.* [131], defines a nozzle with an additional "annular recessed groove" around the nozzle hole, in the nozzle face. It is claimed that this will allow the deposition of both wide and thin tracks with a single nozzle, by adjusting the flow rate.

The machining ability of the nozzle manufacturer also limits the nozzle design. Nozzle (C) has the least amount of material between the internal and the external cone. This material will become very thin if either the bore or cone is off-centre and can lead to early nozzle failure, by the tip breaking off (see [132] for an example photo of a brass nozzle breaking off).

Nozzle internal geometry defined by the length Ln and the cone angle β_n, will determine the pressure drop, if the material properties and nozzle diameters are equal, as presented in Section 2.8.8.

The length of Ln also influences how the melt deforms after extrusion, which depends on the flow stability, where the length Ln versus the exit speed determines if the flow is steady [133]. In general, it is however recommended to keep Ln as short a possible to reduce the pressure drop. A length of three to four times the nozzle diameter is recommended by [130].

It is shown that an optimal nozzle angle exists for a given set of extrusion parameters by Sukindar and Mohd Ariffin [134]. The pressure drop will decrease, as the nozzle angle increases, until a certain point where turbulent flow and vortices start to occur. The optimal internal angle is determined as the point just before the pressure starts to increase again [115]. The convergent angle ($2\beta_n$) is found to be equal to 130° by Sukindar and Mohd Ariffin [134], for a PLA inlet flow speed of 3 mm/s and a liquefier wall temperature of 160 °C.

The design of the nozzle is, therefore, not trivial and requires careful consideration.

2.8.7 Pressure Drop with Hagen–Poiseuille Flow

The pressure drop ΔP for non-compressible Newtonian flow across a long pipe is modelled as *Hagen–Poiseuille* flow with Eq. (2.36), where L is the pipe length, $\dot{Q}$ the volumetric flow rate and R the pipe radius.

$$\Delta P = \frac{8\eta L\dot{Q}}{\pi R^4} \tag{2.36}$$

Model assumptions are that the fluid is incompressible, time independent, Newtonian and laminar. The pipe is also assumed to be sufficiently longer than the pipe diameter, which has a constant cross section, so that there is no acceleration of the fluid in the pipe [135].

The steady state flow rate $\dot{Q}$ for an FFF can be determined with Eq. (2.37), where v_e is the exit velocity and the nozzle diameter is D_n.

$$\dot{Q} = \frac{v_e \pi D_n^2}{4} \tag{2.37}$$

The *Hagen–Poiseuille* model was adapted by Comb, Priedeman, and Turley [96] for an FFF extruder with Eq. (2.38), where the nozzle pipe

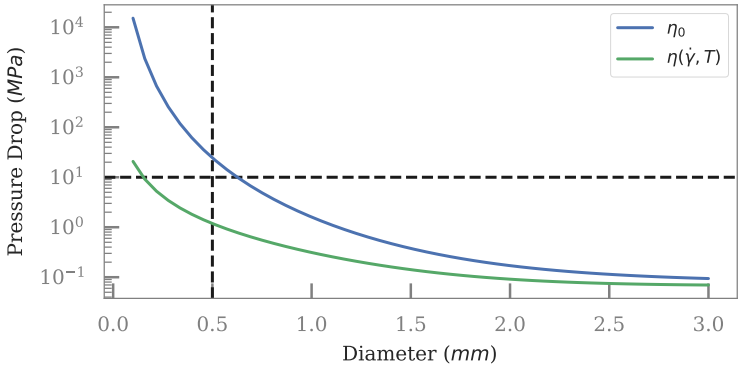


Figure 2.27: Prediction of liquefier pressure drop with Hagen–Poiseuille flow for zero-shear viscosity (η_0) and non-Newtonian, inelastic flow viscosity ($\eta(\dot{\gamma}, T)$), against nozzle diameter at a filament inflow speed of 2 mm/s and a temperature of 215 °C. The dashed horizontal line is the estimated maximum pressure, which the feed mechanism can generate, and the vertical dashed line the current nozzle diameter.

length is Ln, the liquefier length is L_L and the liquefier diameter is D_L. This assumes that there is no tapered section between the nozzle exit diameter and the liquefier diameter.

$$\Delta P = \frac{128\eta\dot{Q}e}{\pi}\left[\frac{Ln}{D_n^4} + \frac{L_L}{D_L^4}\right] \tag{2.38}$$

This model indicates the importance of the nozzle diameter, since numerical analysis finds that the pressure drop along L_L is negligible compared to the drop across Ln. The total pressure drop for nozzles with different diameters are shown in Fig. 2.27, using both η_0 and $\eta(\dot{\gamma}, T)$ at a print speed of 25 mm/s and temperature 215 °C. The maximum pressure, as estimated in Section 2.8.5, is shown as dashed horizontal line and suggests that it should be possible to extrude material at these levels, if the shear thinning viscosity model and parameters (Table 2.8) are used.

The pressure drop over a conical section can also be modelled for a Newtonian fluid with Eq. (2.39) [136], where R_{max} and R_{min} is the entry and exit radius of the conical section with length L, κ the viscosity friction coefficient and ρ the fluid mass density.

$$\Delta P = \frac{\kappa\rho QL}{3\pi\left(R_{max} - R_{min}\right)}\left[\frac{1}{R_{min}^3} - \frac{1}{R_{max}^3}\right] \tag{2.39}$$

The viscosity friction coefficient, κ, is determined with $2\pi\alpha\mu/\rho\,(\alpha - 1)$, where μ the dynamic viscosity and α the correction factor for axial momentum flux. The mean and local fluid speeds are u and $\bar{u}$ [136].

The Hagen–Poiseuille Eq. (2.36) is obtained as R_{max} approaches R_{min} for a parabolic velocity profile ($\alpha = 4/3$) [136].

2.8.8 Pressure Drop for non-Newtonian Fluid Flow

The generalised Reynolds number for a power-law fluid can be used to determine if the flow is laminar or turbulent with Eq. (2.40), where u is the average velocity, D_h is the hydraulic diameter and the flow section geometric properties are a and b.

$$Reg = \frac{\rho {u}^{2-n} D_h^n}{8^{n-1} K \left(b + \frac{a}{n}\right)^n} \tag{2.40}$$

The hydraulic diameter D_h of a circular section is equal to the pipe diameter, while a is equal to $1/4$ and b equal is to $3/4$. A conical section has a hydraulic diameter of $2R\,(1 - \sigma)$, where $\sigma = R/R_o$, with R and R_o the minimum and maximum pipe radii respectively. The values for a and b, for a conical section can be read from a table, for a given σ [112].

The Reynolds numbers for the RF1000 liquefier sections were determined for a print speed of 100 mm/s and found to be below 2100, which indicates that the flow should be laminar, assuming non-compressible, time-independent, power-law fluid flow. The geometric variables (a and b) were estimated as 0.4574 and 0.962 35, while the power-law coefficients, n and K, for the PLA material [100] were assumed to be equal to 0.25 and 41 642 $\mathrm{Pa \cdot s^n}$.

The pressure drop for a power-law fluid can be modelled with Eq. (2.41), if the flow is laminar and by estimating the friction factor f as $16/Re_f$ [112].

$$\Delta P = \frac{2 f L \rho {u}^2}{D_h} \tag{2.41}$$

2.8.9 The Sectioned Liquefier Model

The liquefier pressure drop can be estimated for each of the three liquefier sections. The first section is defined as starting at the melt front and ending at the conical section start, with length L_m (see Fig. 2.23). The second part is the nozzle conical section, with length L_{cn}, and the last section is the nozzle tip, which is another constant diameter pipe, with length L_n.

These three sections as well as the viscosity temperature dependence were modelled by Bellini, Güçeri, and Bertoldi [102], based on the work by Michaeli [137]. In the derivation a flux momentum balance for a power-law fluid is applied to each of the sections to model the pressure drop for

a non-Newtonian, time-independent, incompressible, steady viscous fluid flow. Furthermore, it is assumed that there is no slip at the liquefier wall or feed mechanism.

This model is extended by [103] to include two more sections, since the filament flow direction was turned with 90° as it entered the liquefier in older printers [115]. Note that [102] is still often cited in work related to FFF liquefier dynamics and is therefore also of historical significance in this field.

The relationship between shear stress and rate for a power law fluid is given in Eq. (2.42), with m the flow exponent and ϕ the fluidity material constant.

$$\dot{\gamma} = \phi \tau^m \tag{2.42}$$

Reordering Eq. (2.42) results in Eq. (2.43), which is similar to the shear rate power law Eq. (2.33).

$$\tau = \left(\frac{\dot{\gamma}^{1-m}}{\phi} \right)^{1/m} \tag{2.43}$$

The material constants of the two equations can then be defined in terms of each other by comparison. The flow exponent is expressed as $m = 1/n$ and the fluidity constant as $\phi = K^m$. Substituting these expression into the shear rate model gives the viscosity in terms of ϕ, τ and m with Eq. (2.44) [102].

$$\eta = \phi^{-1} \tau^{1-m} \tag{2.44}$$

The pressure drop ΔP_i for each section (i) can be modelled with Eq. (2.45) [102, 105, 138], where subscript L and N refers to liquefier and nozzle geometry and L, R and D refers to length, radius and diameter. The entry velocity is defined as v and is uniform, since the filament is not

yet melted.

$$\Delta P_1 = 2L_L \left(\frac{\upsilon}{\phi}\right)^{1/m} \left(\frac{m+3}{R_L^{m+1}}\right)^{1/m} \tag{2.45a}$$

$$\Delta P_{2a} = \left(\frac{2m}{3\tan(\beta/2)}\right) \left(\frac{1}{R_N^{3/m}} - \frac{1}{R_L^{3/m}}\right) \tag{2.45b}$$

$$\Delta P_2 = \Delta P_{2a} \left(\frac{\upsilon}{\phi}\right)^{1/m} \left(R_L^2 (m+3)\, 2^{m+3}\right)^{1/m} \tag{2.45c}$$

$$\Delta P_3 = 2L_N \left(\frac{\upsilon}{\phi}\right)^{1/m} \left(\frac{(m+3)\, R_L^2}{R_N^{m+3}}\right)^{1/m} \tag{2.45d}$$

The temperature dependence of the viscosity is modelled with the Arrhenius relation with Eq. (2.29), which can be expressed with Eq. (2.46), where T_α is a reference temperature for which $\eta_0(\dot{\gamma})$ and $H(T_\alpha) = 1$ is determined, α is the activation energy and T_0 is a conversion factor from degrees Celsius to Kelvin ($T_0 = 0$ for absolute temperatures) [102].

$$\eta(T) = H(T)\, \eta_0(\dot{\gamma}) \tag{2.46a}$$

$$H(T) = e^{\left[\alpha\left(\frac{1}{T-T_0} - \frac{1}{T_\alpha - T_0}\right)\right]} \tag{2.46b}$$

The total pressure drop ΔP_{total} as a function of the temperature, power law material constants, inlet speed and liquefier geometry can be determined with Eq. (2.47), by summing each pressure drop section in Eq. (2.45) and multiplying it with $H(T)$ (Eq. (2.46b)), with the additional assumption that the melt is at a uniform temperature T, equal to the exit temperature, with a constant specific heat capacity cp.

$$\Delta P_{total} = H(T) \sum_{i=1}^{3} \Delta P_i \tag{2.47}$$

2.8.10 Equivalent Models

Electronic equivalent circuits have been used to model the liquefier. The volumetric flow rate is taken as the current and Ohm's law is used to model Poiseuille's law.

The RL Extruder Model

A first-order RL circuit for the liquefier is proposed by Bellini, Güçeri, and Bertoldi [102], where the feed mechanism slippage is modelled as a resistance R, while the material properties ($\eta(T, \dot{\gamma})$) and heat flux are modelled with an impedance L.

This circuit analogy is used to derive a transfer function $G(s)$ (see Eq. (2.48)).

$$G(s) = \frac{\mu_{RL}}{T_{RL}s + 1} e^{-\tau_{RL}s} \tag{2.48}$$

The values for the average system gain $\mu_{RL} = 1/R$, T_{RL} (L/R) and the time constant (τ_{RL}) were determined by measuring the system step input response. This was done by increasing the extrusion speed with a Heaviside step function, while printing a track. The track width change was measured after completing the experiment.

The average system gain μ_{RL} is found to be 0.8865. This indicates a steady state error, attributed to the feed mechanism slippage [102]. It is possible to compensate for a constant under extrusion, by adjusting a multiplier parameter in the slicer configuration. The average time delay τ_{RL} is measured as 0.04 s. This time delay, however, cannot be corrected for with conventional slicer settings.

The RC Extruder Model

A method to compensate for step response time constant in real time is patented by Comb [139]. The effect of the increase in the volume of the melt during extrusion is considered to derive an RC equivalent model.

A step change in the extrusion speed will change the volume of molten material in the liquefier. The liquefier pressure will increase, if the volume of material fed into the liquefier (V_i) is greater than the volume flowing out (V_O). This pressure increase causes the exit flow rate to increase, until an equilibrium is reached between the input ($\dot{Q}_i$) and output ($\dot{Q}_i$) volumetric flow rates, after a certain delay time τ_{RC}.

The change in volume ΔV, during this adjustment period for a time step Δt can be written as $\Delta V = V_i - V_O = \left(\dot{Q}_i - \dot{Q}_O\right)\Delta t$.

The patent uses an RC circuit equivalent model, with $C = \Delta v/\Delta\dot{Q}_O$ the capacitance, $R = \Delta P/\Delta\dot{Q}_O$ the flow resistance and the time constant

τ_{RC} equal to RC, to derive the differential equation Eq. (2.49) for the first order liquefier pressure dynamics.

$$\frac{\delta \dot{Q}}{\delta t} = \frac{\dot{Q}_i - \dot{Q}_o}{\tau_{RC}} \tag{2.49}$$

The required input flow rate can then be determined in real time with Eq. (2.50), where A is the flow output area, vxy the XY-print (head) speed and a_{xy} the acceleration.

$$\dot{Q}_i = A\left(vxy + \tau_{RC} axy\right) \tag{2.50}$$

The capacitance (C) is determined with the melt bulk modulus B, with $C = V_L/B$ and V_L the volume of the liquefier. The volume change $V_{p1} - V_{p0}$, given a change in pressure ΔP, is used to determine B with Eq. (2.51), by using a filament material data sheet [139]. The flow resistance (R) was determined experimentally.

$$B = \frac{\Delta P}{\left(V_{p1} - V_{p0}\right)/V_{p0}} \tag{2.51}$$

Interestingly, the time constant τ_{RC} is found to be 0.03 s [139], which is comparable to the time constant τ_{RL} found by [102]. Both models were based on a Stratasys FDM® printer, with ABS filament.

Higher Order Modelling

The applicability of high order modelling was investigated by Heij [126], who measured the feed rate step response. A black-box maximum likelihood model was derived for the liquefier at low extrusion speeds in order to neglect feed mechanism slippage and an ARMAX (Autoregressive Moving Average Model with eXogenous inputs) was used to determine model parameters. This investigation, however, finds that higher order models do not significantly improve the modelling result. A first order model is therefore recommended [126].

The Filament Advance Algorithm

The time delay effect is modelled by the Repetier Firmware using the optional *advance* algorithm, which determines the number of additional steps (S_{adv}) required to build-up the pressure in the liquefier [72], based on the *Mattroberts-Firmware* [140].

The derivation is based on the elastic compression of the filament Eq. (2.52) with E the Young's modulus, ΔL the length of filament pushed into the liquefier and L_0 the distance between the melt front and filament

feed mechanism contact point. The units are converted from millimetres to steps with $S_{cnv,E}$ (see Section 2.4.3).

$$\Delta P = E\frac{\Delta L}{L_0} = E\frac{S_{adv}}{L_0 S_{cnv,E}} \tag{2.52}$$

The pressure drop is further modelled using Bernoulli's principle, by neglecting gravity and using a friction constant (ζ) Eq. (2.53a), where v_p is the exit speed and v_E the feed speed. The equation is rearranged to obtain Eq. (2.53c), where A_r is the ratio between the filament and the nozzle cross section areas. This also defines the constant K_{adv} (see Eq. (2.53b)), which can be set as a firmware parameter.

$$\Delta P = \left(v_p^2 - v_E^2\left(1-\zeta\right)\right)\frac{\rho}{2} \tag{2.53a}$$

$$K_{adv} = \left(A_r^2 - 1 + \zeta\right)\frac{L_E S_{cnv,E}\rho}{2E} \tag{2.53b}$$

$$S_{adv} = K_{adv} v_E^2 \tag{2.53c}$$

A value K_{adv} range of 40 – 70 is found by [72] for a 3 mm filament and a 0.5 mm nozzle diameter. A short calculation for the RF1000 printer configuration, with a PLA Young's modulus of 3500 MPa [100] and ζ equal to one, results in K_adv equal to 5.6. In practise, however, this value is determined through experimentation.

It is also possible to use the Hagen-Poiseuille model for the pressure drop, assuming laminar flow, as argued by Henschke [141]. This results in a linear relationship between S_{adv} and v_E, instead of a quadratic one, with K_{adv} equal to $32\eta L_n/D_n^2$, given that $\dot{Q} = A_n v_E = \pi v_E D_n^2/4$. The extra extruder steps during acceleration will shift the velocity ramp upwards, while decelerating will have the opposite effect, by reducing the number of extruder steps [141].

2.8.11 Liquefier Temperature Control

The required constant heat flux ($\dot{q}$) through the liquefier wall to the melt can be estimated with Eq. (2.54), where T_i is the filament temperature at the liquefier entrance and A is the cross section of the melt [102, 115]. Here the material density is given by ρ, the flow speed at the cross section area (A) by υ and the heat capacity by c_p.

$$\dot{q} = \dot{m}c_p\left(T - T_i\right) = \frac{\rho\upsilon A c_p}{2\pi R_L L_L}\left(T - T_i\right) \tag{2.54}$$

The printer controls the liquefier temperature with a PID control loop. The temperature is sensed by a thermistor attached to the liquefier wall and the resulting change in voltage is translated to temperature by the firmware, usually with a look-up table. The printer adjusts the PWM resistive heater current in order to reach the target temperature.

Various time delays, therefore, come into effect. The most important requirement is that the filament reaches a certain temperature, above the melting point, before reaching the nozzle conical section, to reduce the pressure drop, since the viscosity decreases significantly with increasing temperature.

The constant heat flux model is defined in Eq. (2.54). The minimum length or melt front position is also important and can be estimated with Eq. (2.55) [127], if it is assumed that the liquefier wall is at a constant temperature, due to the relatively high thermal conductivity of liquefier material.

$$\frac{Tm - Tw}{T_i - Tw} = 2 \sum_{n=1}^{\infty} \frac{e^{-\lambda_n^2 z'_m}}{\lambda n J_1 (\lambda n)} \tag{2.55}$$

The melting temperature, Tm, will be reached at the non-dimensional position z'_m, given the wall temperate T_w and the filament entry temperature T_i. The sum requires the computation of the roots of the zero-order Bessel function of the first kind λ_n, as well as first-order Bessel function of the first kind J_1 [127].

The distance, z_m, which is required to raise the filament temperature from T_i to Tm can be determined with the definition of z'_m Eq. (2.56), where α_d is thermal diffusivity, R_f is the filament radius, assumed to be equal to the liquefier radius, and υ the filament speed [115, 127]. Note that the thermal diffusivity is the thermal conductivity divided by density and specific heat capacity ($\alpha = K_{cond}/(\rho c_p)$), at constant pressure.

$$z'_m = \frac{\alpha_d z_m}{\upsilon R_f^2} \tag{2.56}$$

This model was used to predict the maximum feed speed, which can be reached before the liquefier end is reached in Fig. 2.28. The melt front location was solved by using a convergence check ($\Delta S_m < S_{cvg}$) in the summation (S_m), to determine the number of terms (m). A total liquefier length of 22 mm, with a minimum length of 3 mm, were used as input parameters. The figure shows that the length of molten material decreases with increasing speed, with a melting temperature of 170 °C and the liquefier wall temperature equal to 215 °C.

The patent application by Batchelder, Swanson, and Johnson [142] also derives a similar model for the length required to heat the material, as applied to a liquefier with more than one heated section. The sections can be controlled so that the filament core temperature reaches the required

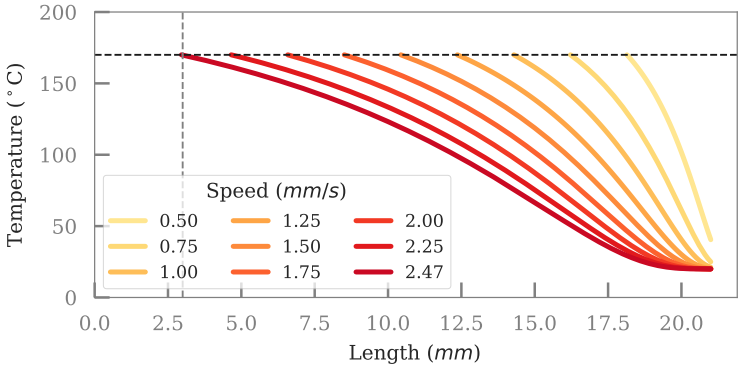


Figure 2.28: Melt front location at different speeds. A maximum speed of 2.47 mm/s is reached before the remaining liquefier length is less than 3 mm (dashed vertical line), with a wall temperature of 215 °C and a melting temperature of 170 °C (dashed horizontal line). The nozzle conical section start is on the left (at 0 mm) and the liquefier entrance on the right (at 22 mm).

temperature, without overheating the filament surface, in order to reach higher extrusion speeds. The end result is a longer liquefier, i.e. for faster flow rates a longer liquefier is required to heat the filament.

Alternative heating approaches include inductive or laser heating. For example, a laser diode is used by Go and Hart [143], along with other modifications, to achieve a 14 times faster extrusion speed.

2.8.12 Feed Mechanism Slippage Modelling

Feed slippage can be categorised in three groups. Firstly, there is the steady state flow error, due to the imperfect force transfer from the feed gear to the filament. Secondly, there is *stick-slip* slippage, where traction is lost, but recovered again. This behaviour is hard to model [126]. Finally, extrusion failure can occur where the feed gear shears the gear teeth impression in the filament off. In this case there is no possibility of recovery, without manual intervention.

Slippage can be reduced by printing at slow speeds, by optimising the liquefier to reduce the pressure drop or by optimising the feed mechanism. The feed mechanism can be improved by increasing the contact area between the filament and the feed gear, optimising the feed mechanism material to improve grip or by implementing closed loop control of the feed mechanism [102, 126, 144, 145].

Furthermore, the gear can also wear during operation, which will also result in a steady state error [126, 144]. This is because the actual

material feed rate depends on the diameter of the feed gear (D_{gear}) and the rotation speed of the motor ω, with $\upsilon = \omega D_{gear}$.

The work of Bellini, Güçeri, and Bertoldi [102] found that the output flow rate has a steady state error, attributed to feed mechanism slippage. The slippage was modelled by the resistance R in the electrical circuit analogy Section 2.8.10, as well as by the generalised Navier law Eq. (2.57), where F_s is the friction force between the feed rollers and the filament, υ_r is the feed roller speed, μ_{f0} and e are material properties taken as 1 and 0.48.

$$F_s = \mu_{f0} \upsilon_r \upsilon_r^{e-1} \tag{2.57}$$

The total force exerted on the melt is determined with $F = F_{max} - F_s$, with F_{max} the force applied by the motor (see Eq. (2.34))).

Slicing algorithms usually account for slippage with a constant feed multiplier. This is determined by extruding a length of filament and measuring the actual length fed into the liquefier. The measured length is compared with the G-code commanded length to determine the correction factor. This assumes that the *steps per millimetre* setting is already set correctly.

Feed mechanism slippage is also considered by Go *et al.* [65], who focussed on the maximum speed and force the feed mechanism can achieve. The filament will fail in shear if the force supplied by the motor F_E exceeds the shear pressure $\tau_s(\upsilon)$, which can be supported by the filament shear area A_s, which is the contact surface area between the filament and the feed gear.

The size of the shear area will depend on the depth of engagement of the gear into the filament, which is determined by the spring force exerted by the pinch bearing on the filament (see Section 2.8.4). This contact area is predicted to decrease with increasing feed speed, further reducing the maximum force [65].

The filament will therefore fail in shear, if the back pressure ΔP of the extruder is greater than the maximum allowable shear force. These relationships, which predict when the filament will fail in shear are given in Eq. (2.58) [65]

$$F_E < \tau_s(\upsilon) A_s < \Delta P A_f \tag{2.58}$$

The shear strength can be approximated with $\tau_s \approx \sigma_{uts}/\sqrt{3}$, where σ_{uts} is the ultimate tensile strength. The shear rate υ dependency should also be considered though [65].

The feed slippage percentage, α_s, between the feed gear and filament is defined by Heij [126] with Eq. (2.59), where ω is the rotation speed of the feed gear with pitch diameter D_f and υ_f is the filament speed.

$$\alpha_s = \frac{\omega D_f/2 - \upsilon_f}{\omega D_f/2} \tag{2.59}$$

Note that the feed slippage percentage represents how much the actual speed differs from the G-code set point speed (v_{SP}), with $v_f = (1 - \alpha_s)\, v_{SP}$. The non-linear relationship between α_s and v_f can be modelled with the black-box "Magic Formula" of Pacejka, which was developed for the automotive tyre industry [126].

This model can be applied to an extruder by replacing the force term of the original formula with the feed speed, assuming that $\Delta P \propto v_f$ and $F \propto v_F$, which gives Eq. (2.60), where B_s,C_s,D_s and E_s are fitting constants [126].

$$v = D_s \left(C_s \left[B_s \left(1 - E_s\right) \alpha_s + \frac{E_s}{B_s} \tan^{-1} \left(B_s \alpha_s \right) \right] \right) \tag{2.60}$$

2.8.13 Extruder Modelling Discussion

Most researchers use Finite Element Simulation (FEM) to model the complex liquefier dynamics, see for example [43, 101–104, 134]. FEM simulations are, however, difficult to implement in real time control, which motivates the use of equivalent models, such as the RC or RL electrical circuit analogies or the K_{adv} factor [102, 126].

Both the equivalent and analytical models, however, require a method to measure the predicted outcome as well as the relevant process inputs. Current in-process measurement methods are reviewed in Section 2.9.

These models and measurements can the be used for process monitoring, active control or optimisation, where optimisation is used to fine tune an extruder, as discussed in Section 2.10.

The question is what should be optimised. For example, print speed is increased by reducing the resolution and increasing the heat transfer to the filament, which usually requires a longer liquefier. This is limited, though, by material thermal degradation [142]. On the other hand, resolution is increased by reducing the nozzle diameter, which increases the required feed force and the risk of feed mechanism slippage. Furthermore, the reduction of feed slippage increases the final part volumetric accuracy.

The selected extrusion material will also interface with extruder construction materials. Care needs to be taken for special application (for e.g. food safety) and component wear.

2.8.14 Post-Extrusion Modelling

This section focussed on the volumetric flow rate and parameters affecting it. The final part geometrical accuracy and mechanical properties will, however, depend on additional factors, such as the solidification processes, melt viscoelastic properties, slicer path plan, welding of the newly deposited layer with the previous layer, shrinkage and residual stress.

The post-extrusion track shape geometry is simulated by Comminal *et al.* [146]. This is important since the track shape influences the final part mechanical properties and geometry, which is determined by factors such

as the inter-layer bond formation and surface roughness. The slicer also uses a model for the track width to determine the required volumetric flow rate [146].

The cross section track area (A_t) can be determined with volumetric conservation (see Eq. (2.61)), given the nozzle diameter D_n, average nozzle flow rate υ_n and average track deposition rate υ_p [146].

$$A_t = \dot{Q}_n / \upsilon_p = \frac{\pi D_n^2 \upsilon_n}{4 \upsilon_p} \tag{2.61}$$

This assumes non-compressible flow and that the volumetric flow rate of the filament ($\dot{Q}_f$) into the liquefier is assumed equal to volumetric exit flow rate at the nozzle ($\dot{Q}_n$), as well as the volumetric deposition rate ($\dot{Q}_p$).

Interestingly, the cross section track area (A_t) does not depend on the gap (layer) height h_z. This height is, however, used to determine the cross-section width W, by assuming that the track conforms to a certain shape. Three ideal cross section shapes used in literature for the deposited track are shown in Fig. 2.29 [146].

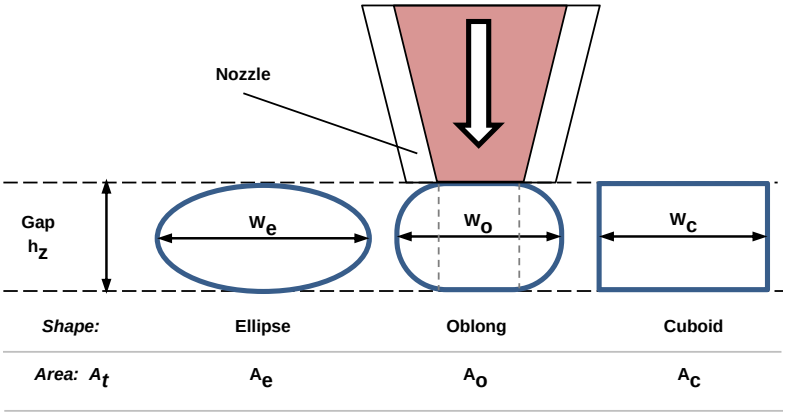


Figure 2.29: Three ideal track cross section shapes, as used in literature, to model the contact surface area for mechanical and geometrical modelling [146]. The slicer also uses a cross section model to determine the track width.

The cross section area (A) and width (W) in terms of υ_n, υ_p, D_n and h_z for each ideal shape is given in Table 2.10, where $\upsilon_r = \upsilon_n / \upsilon_p$ and $H_r = D_n^2 / h_z$.

Table 2.10: Ideal track shapes, areas and widths [146], where the gap height (layer height), flow ratio and the nozzle to diameter ratio are given by h_z, $v_r = v_n/v_p$ and $H_r = D_n^2/h_z$ for shape i.

Shape	**Area**	**Width**
Ellipse	$\frac{\pi}{4} h_z W_e$	$v_r H_r$
Oblong	$h_z W_o - h_z^2 \left(1 - \frac{\pi}{4}\right)$	$\frac{\pi v_r H_r}{4} + h_z \left(1 - \frac{\pi}{4}\right)$
Cuboid	$h_z W_c$	$\frac{\pi v_r H_r}{4}$

The track area can be written in terms of filament feed rate and diameter with $A_t = \pi D_f^2 v_f / 4 v_p$. The slicer used in this work uses the *oblong* shape, along with the layer height and a pre-set width to determine the amount of material required to deposit the track [42].

The width is determined by the position of the track within the object. An external parameter, for example, can be printed using a thinner width for increased detail, while an infill track (inside the object) can be printed with larger track widths [42].

The length of filament to feed into the liquefier (ΔE_f) can be determined using Eq. (2.62), where L_{xy} is the track length in the XY plane.

$$\Delta E_f = \frac{4 L_{xy} A_t}{\pi D_f^2} \tag{2.62}$$

The expected track width for oblong shape W_o can also be written with Eq. (2.63). This can be used to predict the ideal track width, if the feed parameters are known.

$$W_o = \frac{\pi \Delta E_f D_f^2}{4 h_z L_{xy}} + h_z \left(1 - \frac{\pi}{4}\right) \tag{2.63}$$

Simulations were used by Comminal *et al.* [146] to predict extruded track shape. The simulation model is based on the assumptions of an isothermal Newtonian fluid and a creeping laminar flow. Significantly, the simulations found that the appropriate ideal shape ranges from the cylindrical ellipsoid form to the flat, rounded edge, cuboid (rectangular) shape, if printing speed is increased and the vertical gap decreased.

The simulations also tested the assumption that the gap height, between the nozzle and build, is equal to the track height. It was found that

the track height varies from −10 % to 35 %. It is also noted by Comminal *et al.* [146] that the nozzle tip shape will affect the track geometry, as discussed in Section 2.8.6.

2.9 Process Monitoring and Closed Loop Control

Process models derived from in-process monitoring data can be used to improve the slicer output or the printer design. Real time process monitoring can also be used for early fault detection and quality assurance. Closed loop control, however, can avoid fault states all together as well as improve printing accuracy and speed.

Examples of printer variables which can be monitored are the position of each axis, feed mechanism slippage, exit flow rate and the material feedstock quality. After deposition, the formation of gaps, voids, under and over extrusion as well as the layer height, can also be monitored.

Closed loop control of the positioning system should improve the accuracy of the final part. For example, a 75 % trajectory error reduction was achieved by implementing closed loop control on a commercial printer gantry by Weiss, Storti, and Ganter [79].

The extruder, which has a nozzle diameter in the range of 0.1 – 1.0 mm, has a far greater impact on the part resolution, than the head positioning, since the microstep resolution of the stages is around 10 µm.

2.9.1 Liquefier Monitoring and Control

The pressure drop models (see Section 2.8.10) motivate the use of closed loop control of the feed mechanism [102].

Closed loop control of the feed mechanism was implemented by Heij [126], who use a rotary encoder to measure the filament feed speed. The feed slippage percentage was determined with Eq. (2.59), by assuming that the stepper motor speed is correct. The measured slippage was then used to control the feed mechanism speed.

An optical microscope was used by Greeff and Schilling [145] to estimate both the gear and filament speed as well as the filament width, to obtain the volumetric flow rate. This method and results are presented in Section 3.1. The method was further developed to measure the offline (not during printing) exit flow (see Section 3.2). The monitoring capabilities were also extended by converting the built-in bed level sensor to a liquefier pressure, for real-time monitoring (see Section 3.2.3).

The offline extrusion pressure was measured by Go *et al.* [65] to determine the rate limits of a FFF printer. The extruder test bench developed by Hofstätter *et al.* [106] also measures the liquefier pressure, in order to develop an FEM model of the liquefier. Both of these works used a strain gauged cantilever beam (referred to as a load cell) to measure the force, which was used to calculate the pressure, with the filament diameter $\Delta P = F/A$.

Stratasys, the company which introduced FDM®, is also active in the field of process control. A volumetric flow controller is patented by

Zinniel and Batchelder [147], which proposes various effective filament area sensors and filament feed speed tachometers. The need to monitor both the filament diameter and the feed speed (slippage) is made clear in the patent.

Filament feed control with velocimetry is described in another Stratasys patent by Batchelder [148]. The laser based velocimetry instrument is incorporated into the nozzle. This is, to the author's knowledge at the time of writing, the only method proposed so far, which can measure the nozzle exit flow rate in real time, during printing.

A special liquefier design, with differentially heated sections, and integrated pressure sensing for feedback control is described in a patent application by Batchelder, Swanson, and Johnson [142]. This also contains an interesting discussion on the melt front development during extrusion.

A vision based system was used by Hoelzle, Alleyne, and Wagoner Johnson [149] to measure the deposited track volume, in order to adjust the flow rate for each layer for a syringe type extruder, while a 2D laser triangulation sensor was used by Faes *et al.* [150] to monitor and control the printed track z-height.

Thermal vision sensing was used by Pollard *et al.* [151], where an IR camera and thermistors are used to monitor the extruder thermal fluctuations during step changes in the feed rate.

On-line melt monitoring should also be possible with conventional methods used for industrial extrusion processes. These methods include spectroscopic, scattering, conductivity and viscosity measurements, as presented by Alig, Steinhoff, and Lellinger [152]. An interesting idea is to measure the melt conductivity, which can be useful for the extrusion of conductive tracks. Furthermore, ultra-sound can be used to measure the FFF liquefier melt properties, but implementation can be challenging due to cost, size and weight restrictions.

2.9.2 Filament and Printer State Monitoring

Real time filament presence and water content measurement with capacitive sensing is patented by Paul and Batchelder [153]. Many filament materials are hygroscopic, requiring special care to reduce water absorption during storage. Water content in the filament may cause air pockets, bubbles, or other detrimental effects to the finished part. This aspect is, however, not yet fully considered in the literature.

Filament presence sensors, which simply provide a digital level by measuring the filament feed speed, can now be bought for FFF printers. The sensor uses an optical rotatory encoder. The resulting signal can be integrated into the firmware or the printer control server, to stop the print when a filament jam or loss of filament supply are detected. This can save a print, but still requires manual intervention [154].

An alternative control concept is presented by Kim *et al.* [155], where the motor current is used as a feedback signal to detect the deposition state, assuming the motor torque will increase due to an increase in liquefier back pressure.

Nozzle clogging was detected by Tlegenov, Wong, and Hong [156] with an accelerometer placed next to the extruder. This was used to determine whether the printer is in an error state or not. Accelerometers were also used by Bukkapatnam and Clark [157] to monitor the printer state and acoustic emission was used by Wu, Wang, and Yu [158], to determine if the printer is in an error state.

A heterogeneous sensor array was implemented by Rao *et al.* [159] for online process fault detection. The sensor array includes thermocouples, accelerometers, an infrared (IR) temperature sensor and a video borescope, where the IR temperature sensor was used to monitor the deposited material melt pool. The sensor array was used do detect defects and optimise printer parameters. Another sensor array approach was presented by Baumann *et al.* [160], which can in principle interface with acceleration, temperature, orientation and hygrometer sensors.

2.9.3 Deposited Layer Monitoring

A microscope camera system was used by Fang *et al.* [161] to capture each layer after completion. Image processing was then used to detect extrusion errors in the photographed layers. The microscope was mounted on the print head and can be scanned across the whole print area. This requires additional time per layer. It was used to monitor the build and gave insight into the process, which can be used to optimise it.

A 3D vision based, part error detection technique is described in Baumann and Roller [162] as well as by Straub [163], where a stationary camera monitors the build, which enables remote monitoring and early fault detection.

Dual camera, real-time build monitoring was used by Nuchitprasitchai, Roggemann, and Pearce [164], which provides a 360° view of the print and was used to detect both failure and normal printing states. This technique uses low-cost web cameras, which makes it applicable to FFF type printers.

2.9.4 Monitoring and Control Discussion

Monitoring the process input variables provides the ability to improve the accuracy. For example, by reducing the amount of slippage the under-extrusion will be reduced, improving the final part [145]. Implementing real time control without significantly increasing the printer cost is, however, a challenge. See, for example, the work on implementing closed loop control on the positioning system [79], where a dedicated controller for each axis is required. This also increases system complexity considerably. Many researchers therefore use monitoring only for early fault detection and off-line process optimisation.

The ability to monitor the object layer by layer provides measurements which can be used for real time process optimisation or correction. This raises the need of heal or repair strategies, which will most likely be of interest in the near future for all AM processes.

2.10 Process Optimisation

One of the first works on 3D printing optimisation was probably presented by Comb, Priedeman, and Turley [96], presented on conference proceedings as work done for Stratasys. The paper reviews the effect of material properties and processing temperatures on the ability of the extruder to process the material. Interestingly, Design of Experiments (DOE) is used to reduce the number of experiments required to understand the relationship between material properties and processing temperature.

The printing process can be optimised by improving each sub-module, for example the build rate was optimised by Go and Hart [143], after they determined the rate limits of all the sub-modules [65].

2.10.1 Parameter Optimisation

Alternatively, the process parameters can be optimised. Both of these methods require a model in order to predict an optimal operation point in the design space. Deriving and validating the model can be done through physical analysis, simulation and experimentation. Experimentation requires measuring the desired response, the process output, and comparing it with the measured inputs or factors.

The number of experiments required to model individual factor responses as well as the interaction effects can be reduced with DOE. Note that there are a lot of parameters, since FFF is a complex process due to the thermal and speed dependency of the filament and the three-dimensional nature of the part. Finally, the desired outcome can also be complex, as discussed in the introduction of Section 2.8.

The following responses are frequently seen: the final part (1) mechanical properties (2) geometric (dimensional) tolerance or (3) the cost (built time, energy or material use). These responses are quantitative. Qualitative responses, however, are also used. They are for example voids, warping or the surface condition [165].

The combination of the different responses into a unified cost model is also a research topic. For example, Peng, Xiao, and Yue [166] optimise the dimensional error, warp deformation and build time with a fuzzy inference system to form a single comprehensive response.

Note that many of the responses are also inter-related. For example, reducing a qualitative response such as internal voids will affect quantitative part strength response.

Analytical modelling of the strength of the printed parts involves predicting the bonding between deposited tracks and layers, which depend on slicer settings such as the build orientation, infill and the temperature. See for example the review by Mohamed, Masood, and Bhowmik [167], simulations by Xia and Lu [168] and experimentation by Li *et al.* [113].

The review by Mohamed, Masood, and Bhowmik [167] fount at least 21 articles, which investigated the effect of build parameters on the printed part, with different DOE approaches, including factorial designs, Taguchi's method, Response Surface Methodology (RSM) and fuzzy logic. Finding the optimal set of parameters for FFF is therefore an active research topic

and also an interesting one, since the same test object can be printed with different sets of parameter values, for a relatively low cost.

Optimal parameters can improve part mechanical characteristics as well as part geometrical and dimensional precision [169]. For example, Li *et al.* [113] find that the parameters directly affect the tensile strength.

A model for the quality evaluation of the print result is developed by Sun *et al.* [165], which combines both quantitative measurands, such as dimensional accuracy, as well the qualitative aspects, such as infill voids. This model uses both slicer parameters and process variables measured on-line as factors.

The need, however, to determine the optimal set of parameters for such a vast array of factors, for a specific goal, which can also vary between print objects, creates a further challenge. Various complex statical methods based on DOE are therefore intensively used. Selection of the appropriate DOE method is also important, since printing several experiments can be time-consuming. Nevertheless, the results should be reproducible and usable by the rest of the community.

2.10.2 Concurrent Parameter Optimisation

Several print runs to complete all the experiments take a significant amount of time, since each object requires certain pre-processing and post-processing steps, for example, axes homing, heating, build bed preparation and extruder initialisation.

The total printing time can be reduced with concurrent or sequential object printing.

Concurrent printing is where several objects are printed at the same time, layer by layer. A geometric accuracy prediction model is verified by Boschetto and Bottini [117], by using DOE and by printing several objects concurrently, at different build orientation angles. Concurrent printing, however, has the drawback that interlayer effects cannot be gauged.

Single run sequential printing is also possible with FFF, if there is enough vertical clearance. This means that test objects can be printed one after the other in a single print run, as long as they are shorter than the vertical clearance limit.

An opportunity therefore exists to combine DOE with single run sequential printing. This idea is developed and tested in Section 3.3.

3 Extrusion Monitoring, Control and Optimisation

This chapter details work done to improve material extrusion with Fused Filament Fabrication (FFF), where the flow rate is the focus point. The review chapter found that the extruder is complex to model, but critical to the final product.

The design and test results of an optical system, which can measure the volumetric liquefier input flow rate, is discussed first. This includes the results of using the method to perform closed loop control of the feed mechanism. Important results here include the characterisation of slippage at different temperatures and feed speeds, with various PLA colours.

This method is further expanded in the next section to investigate the exit flow rate, by developing a unique marking system in order to optically track the extruded material. The built-in bed level sensor is also converted in this section to monitor the force applied by the feed mechanism, and thereby the melt pressure, in real time. Filament retraction, an important technique used to stop the extrusion, is studied with the exit flow measurements.

Single Print Optimisation (SPO) is introduced next, where the parameters for both the steady state flow rate (studied in Section 3.1) and the retraction (studied in Section 3.2) are optimised in a single print, without using the in-processing monitoring capabilities.

The in-process sensor data were, however, collected during the SPO experiments and this is presented in the last section (Section 3.4). This section combines the post-process and in-process measurement data as well as the liquefier pressure drop model (Section 2.8.9) and material model (Section 2.7.4) in a final discussion on the improvement of FFF extruder.

3.1 Closed Loop Control of the Extruder Feed Mechanism

This section is based on the work presented by Greeff and Schilling [145]. Modifications to the image processing and the software application have been successfully implemented since the publication. These are also included here and expand on the original work. The improvements include faster real-time processing and increasing the camera frame rate from 15 FPS (frames per second) to 30 FPS.

The review of the control and motoring strategies found that the feed slippage is a critical parameter (see Section 2.9). The method presented here allows for the volumetric flow rate to be estimated, by measuring the filament speed and width, where the measured width is assumed to be equal to the filament diameter.

The speed of the filament feed gear is also measured. This allows for the feed slippage to be measured, independent of the commanded feed speed, which makes control easier, but also reduces the uncertainty introduced by assuming that the gear speed is equal to the commanded speed. This differentiates this method from the optical encoder based method, as presented by Heij [126]. The ability to measure the filament width differentiates it further, and allows for the monitoring of the feedstock dimensional quality.

The implementation of this method is cost effective. Only a relatively inexpensive digital microscope is required. All the software were developed in Python, which is free and open source. Limitations of this method are also considered, which includes an uncertainty estimation and a prediction of the maximum measurable speed.

The measurement method is discussed first, which includes the mechanical mounting, speed estimation algorithm and the filament width detection. Measurement uncertainty is discussed next, followed by the results of free-air extrusion tests, which characterise the feed slippage. Finally, the measurement method is used to reduce slippage of a special test object, with closed loop control of the feed mechanism.

3.1.1 Vision System Mounting

The Renkforce RF1000 feed mechanism, as presented in Section 2.8.4, can be easily monitored with a USB microscope camera.

The microscope camera mounting on the RF1000 X-axis carriage is shown in Fig. 3.1. All the mounts were printed with the RF1000 printer. A cable guide from *Thingiverse*, called *Printrbot Simple Metal Y-axis Cable Holder* (Thing:2347681) by Trey Fortmuller, was used to safely guide the camera USB cable.

The camera mount was specifically designed for this project and allows for fine vertical positioning of the camera with two screws and inserted nuts. The CAD of the mount is shown in Fig. 3.2, with a few dimensions for scale.

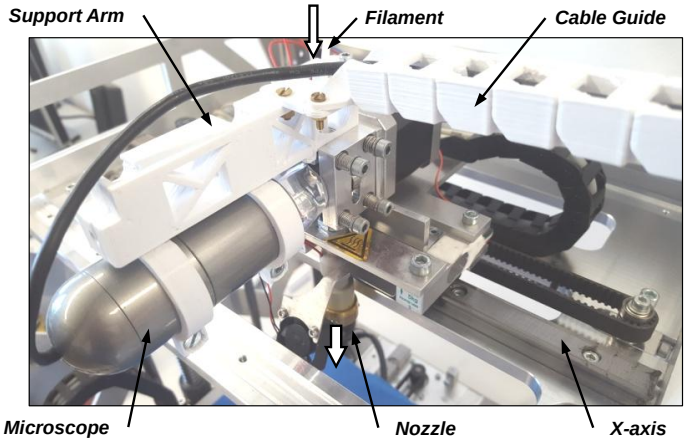


Figure 3.1: Microscope camera mounting on the RF1000, with the printed mounts, including the support arm and cable guide. Large arrows indicate the material flow direction (downwards in image).

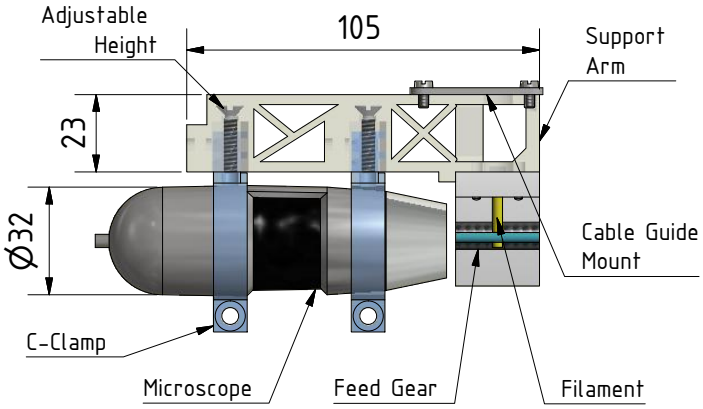


Figure 3.2: Side view of the microscope camera mount design, with height position adjustment screws.

The camera is fixed to the RF1000 carriage by the support arm, which is mounted to the RF1000 bracket, which in-turn is fixed to the extruder stepper motor mounting plate. Two additional holes were drilled into this bracket in order to fix the camera mount. The pinch bearing is also connected to this bracket with tensioning screws (see Section 2.8.4).

The vertical liquefier movement, due to the force applied to it by the filament, does not influence the camera position, since the camera is directly fixed to the X-axis carriage. Vibration or shaking of the camera also appeared to be negligible during printing of objects with short and fast X-axis moves. It is therefore assumed that the camera position is fixed and does not change during tests.

3.1.2 Measurement Method

An image processing algorithm was developed to measure the speed of the gear and the filament as well as the filament width. An application was also written, which could interface between the print server and image processing, allowing for real time printer control and monitoring. Communicating to the printer is achieved through the Repetier-Server [170] with a Websocket interface [171].

A multi-processing (i.e, for running an application on a different processor) *pipe* is used to connect the main application with the image processing program. The main application can therefore be used to perform automated experiments (direct execution of G-code commands) or to print whole objects, while recording the reported temperature as well as the speed and width measurements in real time.

The machine vision algorithm is implemented using the OpenCV library [172] and the Python programming language. Sequential video frames are captured by the camera and analysed in real time. A video file can also be recorded and used for off-line analyses.

Each frame is first pre-processed, after which three Regions of Interest (ROIs) are separately analysed. These image regions, shown in Fig. 3.3 are (1) the area over the gear, (2) the unprocessed filament width above the gear contact point and (3) the filament after (below) the gear contact point. The gear ROI is used to measure the gear speed, while the pre-gear ROI is used to estimate the filament width. The filament speed into the liquefier is measured in the last ROI. This ROI can also be used to analyse the gear tooth profile, which is pressed into the filament.

A black background plate was printed and installed on one of the pinch bearing screws. This plate improves the contrast and feature detection, by hiding the reflective metal bracket in the image background (see Fig. 3.3). The inset image shows the background plate with a blue cover, but it was found that the black cover performed better. It was possible to measure white, green and clear PLA filament using the blue and black covers. The plate colour can, however, be changed to another colour, if the filament colour is, for example, black. The important aspects of the cover are that it is diffuse, does not reflect the light from the camera LEDs and does not interfere with the feed mechanism.

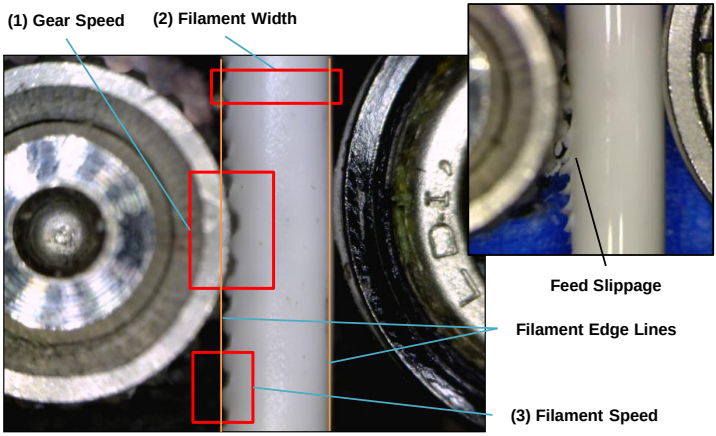


Figure 3.3: Regions of Interest used to measure (1) the feed gear speed, (2) filament width and (3) the speed. The vertical lines are the nominal filament edges and the inset shows feed mechanism slippage.

3.1.3 Speed Measurement with Optical Flow

The speed measurement is realised by using the Lukas-Kanade optical flow algorithm, with image pyramids [173, 174]. Different optical flow methods available in the standard OpenCV library were experimented with, but this specific method was found to work the best.

The principle of the optical flow algorithm used here is as follows. Two consecutive frames are compared to determine the optical flow between them. Points of interest, so called *strong corners*, are identified in the first frame. These points are then searched for in the second frame. Points, which are found again, are used to determine the average point displacement, which is used to resemble the movement, or optical flow, between the two frames. This is extended to several frames by updating searching for new points to track at a set interval (for example every fifth frame).

The result is a point track history for each strong-corner. The point speed can be determined by the relative displacement between two points in a track, multiplied by the frame rate. The track history of a certain point is shown in Fig. 3.4, for three filament speed ROIs, five frames apart, where each ROI is first preprocessed. The circles indicate the detected positions over time, for the example point $P_{j,t}$ at time t.

The current and previous coordinates of each point track are used to determine an average displacement between the given frame in pixels. The speed in pixel per frame is converted to pixel per second using a constant frame rate (FPS) and to millimetre per second using a linear calibration of the image.

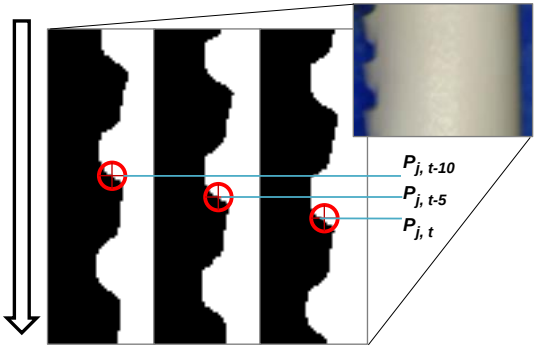


Figure 3.4: Example of the optical flow of a single point $P_{j,t}$, which generates a point track history. The last two points in the track is used to determine the point speed.

3.1.4 The Mathematical Optical Flow Model

Optical flow methods are based on three assumptions: brightness consistency, temporal regularity and spatial consistency.

It is assumed that the tracked pixel brightness does not change over the range of motion. An example of brightness *inconsistency* would be shadows. Secondly, the time between frames, relative to the pixel speed, is slow enough to use differentials as the displacement is small. Thirdly, it is assumed that neighbouring pixels have similar motion. This requires that pixel motion is in the same plane [175]. Note that the third assumption is a special case used to solve the resulting model found with the first two assumptions, as discussed next.

The brightness consistency assumption is defined with Eq. (3.1), where $I\,(x, y, t)$ is the pixel intensity and Δx and Δy is the distance the pixel moved during time Δt.

$$I\,(x, y, t) = I(x + \delta x, y + \delta y, t + \delta t) \tag{3.1}$$

A Taylor series expansion of the brightness consistency assumption, defines the temporal regularity assumption with Eq. (3.2), which is valid for small displacements.

$$I(x + \delta x, y + \delta y, t + \delta t) \approx I\,(x, y, t) + \frac{\delta I}{\delta x}\Delta x + \frac{\delta I}{\delta y}\Delta y + \frac{\delta I}{\delta t}\Delta t \tag{3.2}$$

The general optical flow equation Eq. (3.3) can then be derived by combining Eq. (3.1) and Eq. (3.2), since $(\delta I/\delta x)\Delta x + (\delta I/\delta y)\Delta y =$

$-(\delta I/\delta t)\Delta t$, where the image derivatives are determined with $\delta I/\delta x$, $\delta I/\delta y$ and $\delta I/\delta t$ at (x, y, t) in the corresponding direction [176].

$$\bigtriangledown I^T \mathbf{v} = -I_t \tag{3.3}$$

The motion vector is defined as $\mathbf{v} = [u, v]$ and contains the two optical flow variables, $u = \Delta x/\Delta t$ and $v = \Delta x/\Delta t$. The optical flow is determined by solving $\mathbf{v}$.

For a single pixel this equation is, however, under-determined and various assumptions are made to solve it [176]. The spatial consistency assumption is used by the Lukas-Kanade method [173]. It groups neighbouring pixels together to obtain an over-determined system, which can be solved with the least squares technique, according to Eq. (3.4), with $A\mathbf{v} = \mathbf{b}$, where $\mathbf{A}$ and $\mathbf{b}$ are the stacked displacement and time derivatives [175].

$$\mathbf{v} = \left(\mathbf{A}^T \mathbf{A}\right)^{-1} \mathbf{A}^T \mathbf{b} \tag{3.4}$$

The determination of which points are strong corners is critical. Here the OpenCV method *goodFeaturestoTrack* is used and results in an invertible matrix $\mathbf{A}$. This provides a sparse collection of points to search for with the optical flow algorithm. Dense methods can also be used to compute the optical flow tracking points or the optical flow itself, but is more computationally expensive.

The optical flow equation only applies for small displacements. Larger motion tracking is achieved by using hierarchical image pyramids, where the flow is determined from coarse to fine resolution versions of the frame, by a pre-set number of steps.

3.1.5 Optical Flow Code Implementation

The implementation of the iterative point selection and tracking methods used in this work is based on the examples found in [175, 177]. The OpenCV tutorial example [177] includes a backward-flow calculation (current frame to previous) to filter the tracked points.

The first code snippet (Listing 3.1) determines the value of the variable *speed*, which is the filtered mean optical flow speed of all the tracked points, between the previous frame (*prev_gray*) and the current frame (*gray*).

The process steps are made clear in the examples by using an iterative *for* loop. In this work the computation time was improved by only estimating the optical flow in the pre-set ROIs and by replacing the loop structure in the example with arrays. It is further assumed that the motion is only in the vertical y direction.

Both the forward and reverse optical flows are calculated with *calcOpticalFlowPyrLK*, where two grey scale image frames, the points to be tracked (*p0*) and a constant set of method parameters are provided as arguments. The function returns an array of new points *p1*.

Listing 3.1: Code snippet of the implementation of the optical flow algorithm in Python (the original code is slightly modified to be displayed here).

```
1    p0 = P0.reshape(-1, 1, 2)
2    re_fwd = cv2.calcOpticalFlowPyrLK(prev_gray, gray, p0,
3                                      None, **lk_params)
4    p1, st1, err1 = re_fwd
5    #reverse flow
6    re_bck =  cv2.calcOpticalFlowPyrLK(gray, prev_gray, p1,
7                                       None, **lk_params)
8    p0r, st0, err0 =re_bck
9    d = abs(p0-p0r).reshape(-1, 2).max(-1) #error distance
10   good_d = d < lk_recalc_max_d #select "good" points
11
12   if lk_filter_err:
13       err1[st1 == 0] = lk_err_min + 1
14       error_ok = np.abs(err1) < lk_err_min
15       good_d = np.logical_or(good_d,error_ok.flatten())
16
17   P0 = P0[good_d]
18   P1 = p1.reshape(-1,2)[good_d]
19
20   if P1.size > 0:
21       dy = P1[:,lk_speed_axis] - P0[:,lk_speed_axis]
22       #discard abs(speed) < lk_zero_speed
23       indices = np.abs(dy) > lk_zero_speed
24       speed_xy_per_track = dy[indices]
25
26       if speed_xy_per_track.size > 5:
27           speed = reject_outliers(speed_xy_per_track)
28
29       elif speed_xy_per_track.size  >= 1:
30           speed = np.mean(speed_xy_per_track)
31
32   P0 = P1[indices]
```

The deviation between the original points *p0* and the reversed calculated coordinates *p0r* should ideally be zero and can be used to filter out incorrect points, along with the error values, *err1*, returned by the function.

The displacement in the image Y-axis is determined, after selecting the *good* points. This can be changed to the X-axis, or to an XY distance. Small displacements, due to white noise in the image, are also filtered out in Line 23, based on a fixed threshold level.

The average speed is determined depending on the number of tracked points remaining. Either a simple mean is used or a method called *reject_outliers* further filters the speed. The resulting speed value is reported to the main application, where it can be processed further. The last step in the listing assigns the filtered current points as the previous points, so that it can be used in the next frame.

The next processing step is given in Listing 3.2, which searches for new points suitable for tracking. This is required since the filament or gear

Listing 3.2: Code snippet of the detection of strong corners (the original code is slightly modified to be displayed here).

```
get_new_points = frame_interval >= lk_detect_interval
if P0.size == 0 or get_new_points:
    frame_interval = 0
    pnts_mask = np.zeros_like(gray) + 255
    P0_int = np.int32(P0)
    for x, y in P0_int:
        cv2.circle(pnts_mask, (x, y),
                   pnt_mask_radius, 0, -1)
        new_p = cv2.goodFeaturesToTrack(gray,
                                        mask = pnts_mask,
                                        **feature_params)

    if new_p is not None:
        if lk_refine_subpix: # refine corner locations
            cv2.cornerSubPix(gray,new_p,
                             **subpix_params)
        P0 = np.append(P0,np.float32(new_p.reshape(-1,2)))
        P0 = P0.reshape(-1,2)
```

will continuously be moving across the ROI. The implementation of this code is also based on [177]. The search is triggered if there are currently no points left to track or if a certain number of frames have passed.

First, a mask is made around all the remaining points by drawing circles around them. New *strong corners* are then searched for in the masked image with *goodFeaturesToTrack*, using a predefined set of parameters, *feature_params*.

The location of new points can be refined using *cornerSubPix*, which results in sub-pixel estimation of the corner coordinates. Finally, the new points are added to the existing points and the first code snippet is called again, when a new frame is available.

The whole algorithm is called as soon as a new frame is available from the camera. Threading is therefore used so that the machine vision application can perform other tasks, while it waits for a new frame. A new frame triggers an interrupt, which is handled as soon as possible. The whole machine application is run with the python multi-processor library. The printer controlling main application should therefore run on another processor, further ensuring that frames are not missed or the image processing application overloads (freezes).

3.1.6 Filament Diameter Estimation

The filament width is assumed to be equal to the diameter, i.e. it is assumed that the filament is perfectly round. The width estimation only needs to determine the filament edges in the image, before being imprinted with the gear teeth profile. An interesting simplification can therefore be applied, since the information in only one image axis is of interest. The other axis can therefore be averaged to reduce noise.

Listing 3.3: Code snippet of the filament width measurement method (the original code is slightly modified to be displayed here).

```
1 ave_over_columns = np.average(gray_roi,0) #for gray scale ROI

2 diff_ave = np.diff(ave_over_columns) #first derivative
3 ...
4 # x_lim is either -1 or w_zoom_roi_middle
5 #find array index max/min of first derivative
6 if left_edge:
7     turn_pnt = diff_ave[:x_lim].argmax()
8 else:
9     turn_pnt = diff_ave[x_lim:].argmin() + x_lim
```

The basic width algorithm works as follows. First the colour image is converted to grey scale and preprocessed with some filters (e.g. with a Gaussian kernel blur). The image is then magnified to achieve sub-pixel resolution and sub-ROIs are selected around the approximate location of the filament edges.

The location of each edge is determined in each of the magnified ROIs using the code example given in Listing 3.3. The whole ROI is averaged over the image columns (Line 1), resulting in a single row containing the average edge location.

The assumptions of an approximately vertical edge and that the edge is located within the predefined ROI area, means that the first derivative can be used to find the edge location (Line 2). The left edge is found at the location of the derivative maximum (Line 7) and the right edge at the minimum (Line 9). Note that the methods, *argmax* and *argmin*, return the index of the array element with maximum or minimum value, i.e. the column number in this case.

The integer value found in Listing 3.3 is scaled back, by dividing with magnification value and by adding the appropriate sub-ROI offsets, to determine the location of the edges in pixels. Finally, the difference between the detected edges (right minus left) determines the filament width, which is assumed to be the diameter. The measured width is reported to the main application, which records and filters the results.

Example width measurement results of white PLA with a constant feed speed are shown in Fig. 3.5. The filament manufacturer states that the diameter is equal to (2.85 ± 0.10) mm. The figure shows that the filament is within this tolerance, but it is noted that it is smaller, with a mean of 2.81 mm and a standard deviation of 16 µm, over a time span of 350 s.

Note that this method can be slightly modified to find the gear profile indentation on the filament. This profile can be used to estimate the feed pitch frequency, which can be compared with the gear pitch and used to approximate the feed speed slippage. The profile can also be used to quantify the depth of indentation, which is a measure of the pinch bearing force.

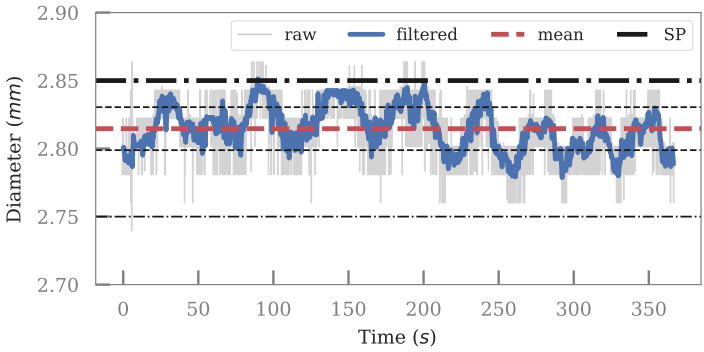


Figure 3.5: Example results of the measured *raw* and *filtered* filament diameter compared to the manufactured supplied tolerances (SP, thick dash-dot), with the mean measured diameter (red thick dashed), standard deviation (thin dashed line) and minimum tolerance level (thin dash-dot)

3.1.7 Speed Conversion and Filtering

Two conversion factors are required, since the image processing returns values in pixels per frame, which needs to be converted to millimetres per second.

The image frames captured by the USB microscope camera are read into program memory with the Python OpenCV library, which represents the frame as a *NumPy* array object. This allows for fast array and matrix operations. Three different USB cameras were successfully tested with this software and any USB camera should work in principle, as long as the appropriate drivers are installed, which makes the work reported here accessible and reproducible by other researchers.

The OpenCV library also has a camera parameter setting interface and the feed mechanism machine vision application uses it to report all camera parameters, which it could read from the camera, in the application GUI (Graphical User Interface). An important parameter is the FPS, which had a default value of 15 FPS. This was changed to 30 FPS, to increase the maximum measurable speed.

The actual frame rate was also estimated by measuring the CPU reported clock interval between frames. The accuracy of this was, however, not clear and not used to update the speed calculation. Note that the value was approximately correct, but that it fluctuated to a degree that it degraded the speed measurement results. A constant frame rate equal to the average value for a specific camera is used instead.

The pixel conversion factor assumes that there is no significant image deformation in the ROI. The filament width was measured three times with a digital micrometer, which has a manufacturer stated uncertainty

Listing 3.4: Code snippet of the speed conversion and filtering.

```
1 gaussian_filter = gaussian(30, 15)
2 gaussian_filter_norm = gaussian_filter/gaussian_filter.sum()
3 speed_conversion_factor = cal_factor*cam_fps
4 speed_mmps = speed_step_per_frame*speed_conversion_factor
5 speed_mmps = convolve1d(speed_mmps,gaussian_filter_norm)
```

of ±2 µm. The micrometer measured width w_μ is then compared to the estimated width in pixels w_{pxl} to determine a linear calibration factor k_{cal} with Eq. (3.5).

$$k_{cal} = \frac{w_\mu}{w_{pxl}} \tag{3.5}$$

The value for k_{cal} in this work is around 0.0197 mm/pxl, this implies a nominal filament diameter of 145 pxl (2.85 mm) and a resolution limit of ±10 µm, without sub-pixel resolution refinement.

The speed in pixel per frame (ΔS_{pxl}) is converted to millimetre per second (υ) by multiplying it with the calibration factor and the camera frame rate, f_{FPS} (see Eq. (3.6)). The speed is also filtered by convolving the speed with a Gaussian window (Wg), which is created with $g(M,\sigma)/\sum g(M,\sigma)$, where M is the number of points in the window and σ is the standard deviation [178].

$$\upsilon = Wg * \left(k_{cal} f_{FPS} \Delta S_{pxl}\right) \tag{3.6}$$

The Python code used to convert a speed array is shown in Listing 3.4. The Gaussian window is created in the first two lines, the speed conversion is performed in Line 4 and the filtering in Line 5, where *cal_factor* is determined with Eq. (3.5) and *convolve1d* is used to convolve the speed array with the filter.

Examples of measured gear and filament speeds are shown in Fig. 3.6, including the unfiltered (raw) and filtered data.

The standard deviation for the unfiltered data was calculated for each section, indicated with the vertical lines in the figure, are given in Table 3.1. The speed difference $\Delta\upsilon = \upsilon_{filament} - \upsilon_{gear}$ or feed slippage is also shown and increases with feed speed.

The standard deviation also increases with speed and raises the question if it resembles measurement variation, actual speed variation or both. This is considered in the next section.

3.1.8 Measurement Uncertainty and Limitations

A major contributor to the measurement uncertainty of the speed is the point displacement measurement between the consecutive frames.

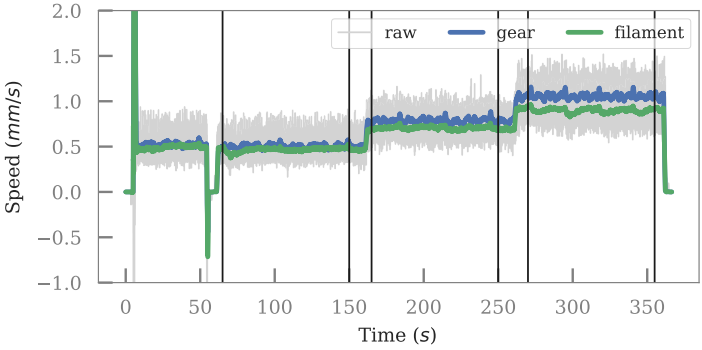


Figure 3.6: Example of measured *raw* and *filtered* speeds. The vertical lines indicate segments of constant feed speed.

Table 3.1: Speed, standard deviation and speed difference for the example measurement sections.

Period	**Gear**		**Filament**		Δv
	Speed	*Stdev.*	*Speed*	*Stdev.*	
s	mm/s				
65 – 150	0.504	0.118	0.465	0.113	−0.038
165 – 250	0.785	0.124	0.706	0.129	−0.079
270 – 355	1.058	0.148	0.900	0.124	−0.158

Furthermore, it is assumed here that the frame time period is fixed and equal to $1/FPS$. The distance measurement depends on the camera calibration and the sub-pixel point detection repeatability, which is also influenced by image noise. It is further assumed that (1) the filament is incompressible, (2) image warping effects over small distances (7 pixels) are negligible and (3) that all the points are in the same plane and parallel to the camera sensor.

The magnification factor (k_{cal}) is around 20 µm/pxl, while a single, ideal microstep will result in a 3.6 µm displacement of the filament (see Section 2.4.3). It is therefore not possible to measure a single microstep with this magnification, but it is also not required here.

The measurement uncertainty is heavily influenced by the pixel to millimetre conversion factor, which depends on the optical and micrometer filament width measurement comparison (see Eq. (3.5)).

The speed uncertainty for a 2 mm/s feed is estimated using the GUM

Workbench® (version 2.4.1.384) [179] software tool, with the model according to Eq. (3.6) and the parameters given in Table 3.2.

Table 3.2: Speed uncertainty parameters for a 2 mm/s feed speed, with abbreviations: Exp. (Expanded), Unc. (Uncertainty), Dist. (Distribution), Rect. (Rectangular) and Cov. Factor (Coverage Factor).

Variable	**Distr.**	**Value**	**Unc.**	**Unit**	**Note**
w_μ	Normal	2850	±2	µm	Exp. Unc. Cov. Factor: 2
w_{pxl}	Rect.	140	±0.5	pxl	Half width limits
ΔS_{pxl}	Normal	3.27	±0.05	pxl	Exp. Unc. Cov. Factor: 1

This finds an expanded uncertainty estimate of (2.00 ± 0.06) mm/s for the speed, with a 95 % confidence interval and a coverage factor of two.

The uncertainty for the micrometer measurement w_μ is based on the manufacturer supplied data, while w_{pxl} is the width estimation without sub-pixel resolution and the ΔS_{pxl} uncertainty is based on the calculated standard deviation.

A speed uncertainty (u_v) of ±0.06 mm/s will result in exit flow uncertainty (u_e)) of $u_v(D_n/D_f)^2$ (with all things else constant), which predicts an uncertainty of the deposited track width of ±20 µm (Eq. (2.63)) for a second-long extrusion. This can be compared to the manufactured specified filament diameter tolerance of ±100 µm, which results in a ±46 µm track width variation. This illustrates how critical the flow speed estimation uncertainty is for derived results, such as the track width.

The frequency content of the speed sections in the measurement example shown in Fig. 3.6 is presented in Fig. 3.7. This was calculated with the *Numpy* discrete Fast Fourier Transform (FFT) implementation.

The peaks shift with increasing speed, for both the gear and the feed speed, indicating a possible relationship between measurement variance and mean speed. This can be due to increased error in the optical flow speed estimation or due to a speed dependant physical effect.

The gear pitch frequency can be considered as a source for this frequency shift, where the feed gear pitch diameter was calculated in Section 2.4.2 with the steps per millimetre conversion factor.

The concept is that the tangential force required by the worm gear to push the filament down varies, since it will cut into (plastic deformation) the filament with every tooth period. This change in force results in a cyclic acceleration, which causes speed fluctuations.

The frequency contents of both the gear and filament speed, at different speed levels, do show a small peak at this frequency (see Fig. 3.6). Other factors should however also be considered, for example aliasing, liquefier

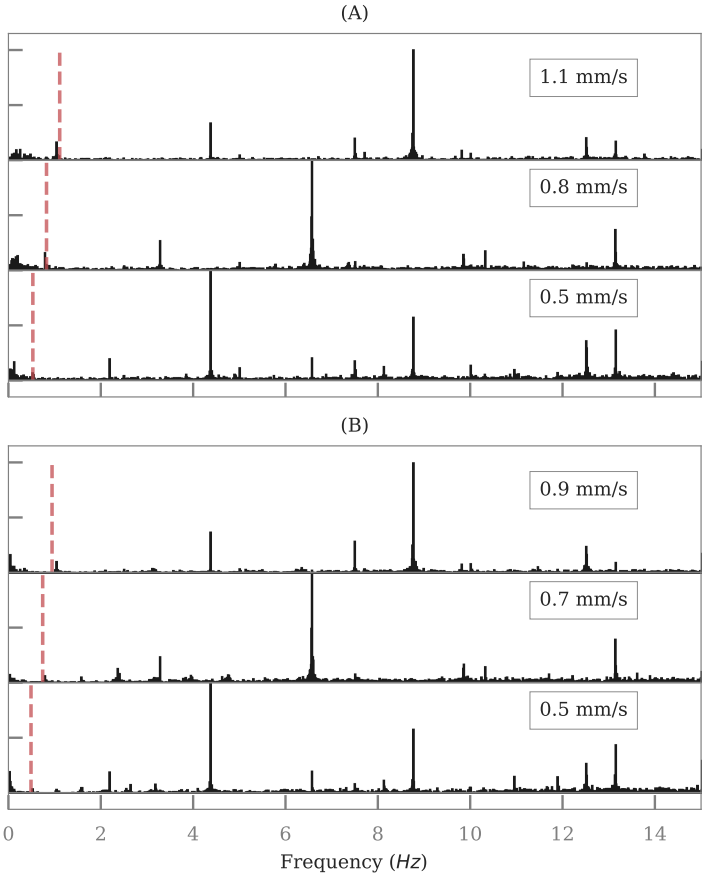


Figure 3.7: Frequency content of the speed measurement example data for (A) the gear speed and (B) the filament speed. Dashed red lines indicate the anticipated gear pitch frequency and text annotate the section mean speed.

back pressure frequency response, stepper motor microstepping and non-ideal motor rotor position response.

The design of a feed mechanism can be aided by considering such data, but this is not the focus of this work. Here, a low-pass filter is applied to smooth the signal, as shown in Fig. 3.8, since the primary interest is the semi-steady state feed slippage.

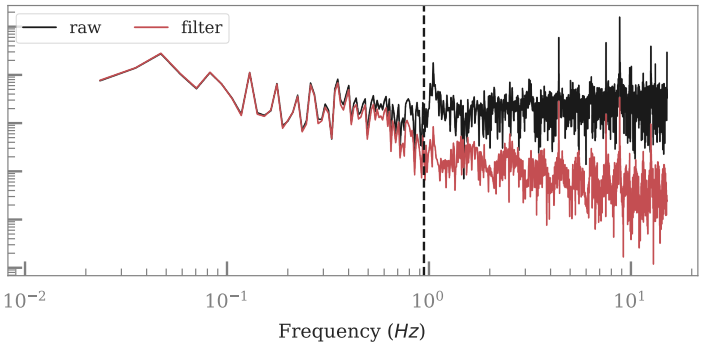


Figure 3.8: Raw and filtered speed magnitude over frequency. The dashed vertical line indicates the gear pitch frequency at the mean speed.

The uncertainty contribution of the speed variation to the mean speed is considered to have a normal distribution. This assumption is checked with Fig. 3.9, where the speed variance ($\Delta v = v - mean(v)$) for the three gear speed sections are shown.

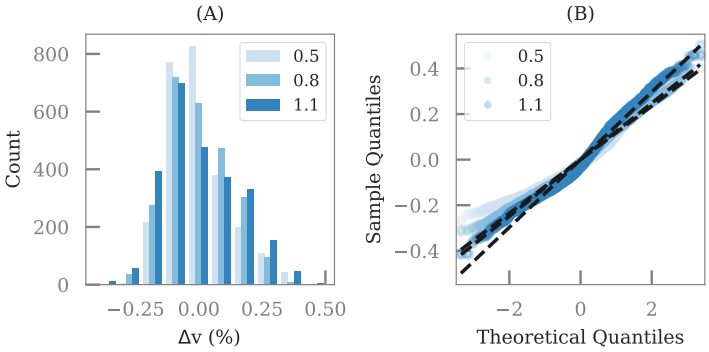


Figure 3.9: Speed variation (A) histogram and (B) quantile-quantile plot for gear speeds 0.5 mm/s, 0.8 mm/s and 1.1 mm/s

The histograms are slightly slanted to the left. This is further investigated with the quantile-quantile (Q-Q) plot. Ideally all the points should be on the diagonal lines for each set, but it is seen to deviate at the lower quantiles. This does not appear to be a significant deviation and it is therefore assumed that the variation has a normal distribution. Interestingly, the Q-Q plot shows that variation becomes more *normal* with an increase in speed.

Another factor is aliasing, which will limit the maximum speed to $v_{max} = k_{cal} f_{FPS} \Delta S_{pxl,max}/2$, with the maximum traceable pixel motion ($\Delta S_{pxl,max}$) assumed to be equal to half of the ROI height.

The ROI height for both the gear and filament are 40 pxl, resulting in a speed limit of 5.9 mm/s at 30 FPS. This is more than double the speed ranges used in this work, except for the fast retraction moves which are around 16 mm/s.

The measurement methods are assumed to be fit for semi-steady state extrusion rates up to 2.5 mm/s, based on the uncertainty estimation presented here. Special care is needed, however, for the derived values such as volumetric flow rate and deposited track width.

Processing Time

Increasing the ROI height will increase maximum measurable speed, but also the computation time required by the optical flow algorithm, while increasing the frame rate decreases the available processing time.

The time required for the speed tracking and width algorithms are compared to the frame period in Fig. 3.10. This shows that the width method requires the most time, while the two speed tracking ROIs require approximately the same time. The variation in the speed tracking is due to the periodical searching for new points to track. A mean of 3.1 ms is used to process all the frames, while a new frame is captured every 33.3 ms.

The maximum frame rate, which can be handled with this processing speed, is 167 FPS (with a safety factor of two), or a maximum speed of 33 mm/s.

3.1.9 Feed Slippage Dependence on Temperature, Speed and Filler

The amount of force required to extrude the material depends on the melt viscosity and the nozzle geometry. The melt viscosity depends on the material, shear rate and temperature. It is expected that feed mechanism slippage will increase, as the back pressure increases, due to increasing viscosity (see Section 2.8.9). This is investigated next, by using the feed speed measurement method described above.

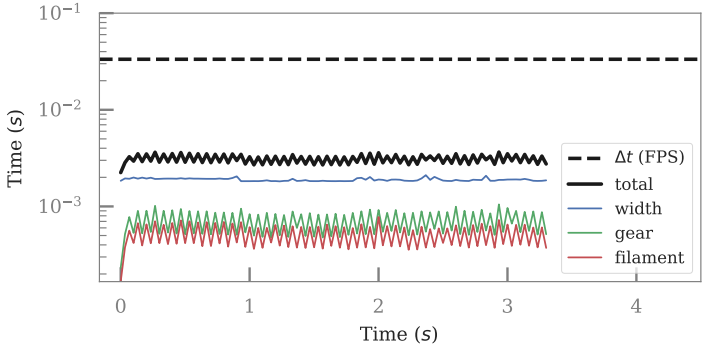


Figure 3.10: Time used to process each ROI (width, gear and filament), as well as the total time. The dashed line is the 30 FPS time period (Δt (FPS)).

The feed slippage percentage Δv_p is calculated with Eq. (3.7), where v_f and v_g are the measured filament and tangential gear speeds, respectively.

$$\Delta v_p = \frac{v_f - v_g}{v_g} \tag{3.7}$$

Three PLA colour filaments were used, namely white, clear and green. The green PLA has a nominal diameter of 3 mm, while the clear and white filaments both have a nominal diameter of 2.85 mm. Test extrusions were performed by sending a **G1** extrusion G-code command to the printer, while the machine vision application monitored the speeds and the filament diameter.

Temperature and Speed Extrusion Tests

Both extrusion temperature and speed were varied during the tests.

The melt temperature was measured with a handheld thermocouple sensor ($\pm(0.5\,^{\circ}\text{C} + 0.3\,\%$ of mv)) and found to agree with temperature reported by the printer. The printer reported temperature is therefore used throughout the rest of this work as the actual temperature.

The manufacturers recommend a liquefier temperature of 215 °C. A testing range of 185 – 230 °C, with steps of 5 °C, was therefore selected, while 50 mm filament was extruded with each test. The tests were repeated three times, if extrusion was possible at the specific test point.

The results of these tests for the white PLA are shown in Fig. 3.11, with box plots for the 2.0 mm/s extrusion speed and dashed lines for the mean speed deviation percentage.

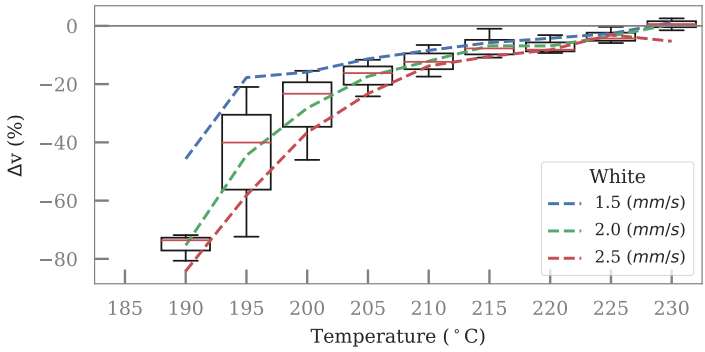


Figure 3.11: Box plot of the extrusion tests with the white PLA at 2.0 mm/s, with dashed lines indicating the mean slippage percentage for each of the three speeds, over temperature.

The feed slippage increases rapidly with temperatures cooler than 200 °C, conversely it decreases at slower extrusion speeds.

The results for all the tests are summarised in Fig. 3.12, where each sub-plot compares the different PLA colour filaments extruded at the same speed. The average mean range (maximum minus minimum) over the three repeat tests, for all the colours and speeds, were 10.1 %. This is equal to ±0.20 mm/s for a 2 mm/s extrusion speed, which is bigger than the uncertainty estimation of ±0.06 mm/s, indicating that other factors influenced the experimental repeatability.

A non-constant filament diameter can, for example, influence the repeatability. The isolation gap between the liquefier wall and the filament will decrease with a larger diameter, which will heat the filament quicker, reducing the melt viscosity [115]. This idea is supported by Fig. 3.12, where the green filament is seen to have the smallest feed slippage percentage.

Filament material has a smaller effect, which is seen by comparing the clear and white filament results at colder extrusion temperatures.

The results confirm the recommended set point of 215 °C, since the slippage temperature dependence becomes approximately constant at this temperature.

The volumetric flow rate error can be estimated by integrating the speed and multiplying it with the measured filament area ($\pi D_f^2/4$). Note that any error in the filament diameter will cause a quadratic error in the volumetric input flow rate.

The standard deviation of the filament width was found to be in the camera resolution range (around 20 µm). The green filament, however, exhibited a larger variation than the clear and white coloured filaments.

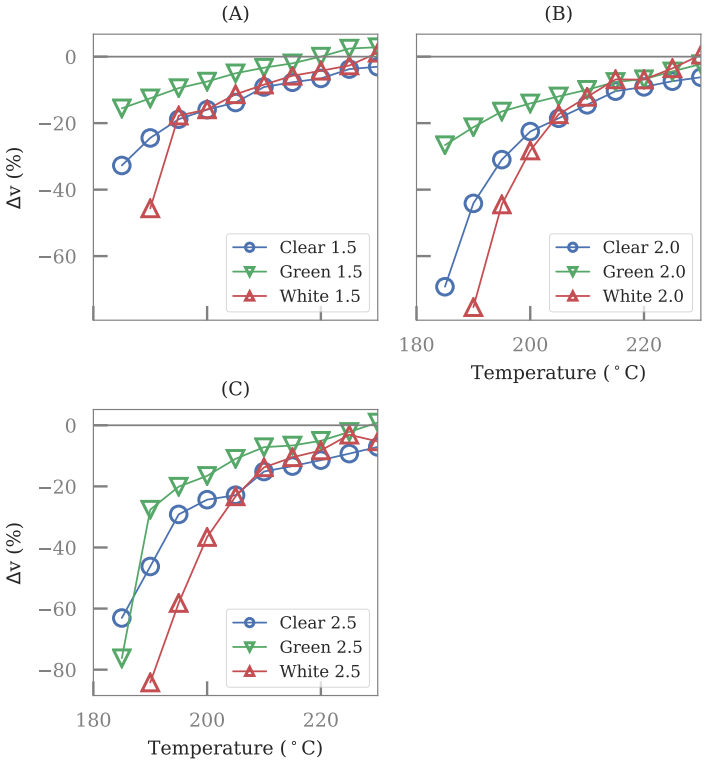


Figure 3.12: Mean feed slippage percentage against temperature for all the tests, at (A) 1.5 mm/s, (B) 2.0 mm/s and (C) 2.5 mm/s extrusion speeds [145].

The volumetric flow rate error percentage will also have the same form as the feed slippage percentage. Implying that under-extrusion was always be present in the test results [145].

Empirical Feed Slippage Model

These results confirm that slippage is dependent on temperature and on extrusion speed. This behaviour can be empirically modelled. The temperature dependence of the viscosity can be represented with an exponential function of the form Eq. (3.8b) (see Eq. (2.29)), where T is the temperature, B is a constant and $A(v_g) = m_a v_g + c_a$ is a linear function

of the feed speed v_g, with constants m_a and c_a.

$$\Delta v_m = \Delta v + 1 \tag{3.8a}$$

$$\Delta v_m = e^{\frac{-A\left(v_g\right)}{T - B}} \tag{3.8b}$$

This model was fitted to the white PLA test results, using a non-linear least squares fit to find $A(v_g)$ and B for each speed, while a least squares polynomial fit was used to find m_a and c_a. Note that the measured speed deviation was first shifted with Eq. (3.8a). The fitting was accomplished using the *SciPy* and *NumPy* functions *curve_fit* and *polyfit* [178].

The fit results are shown in Fig. 3.13 and represents the general feed slippage trend well. This can also be adapted to predict the volumetric flow rate error, given a constant filament diameter.

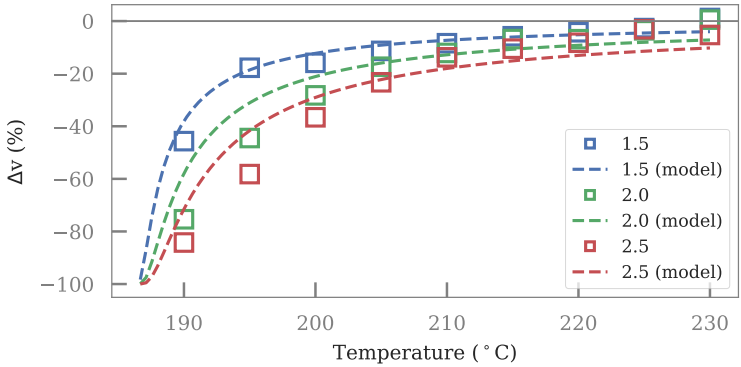


Figure 3.13: Empirical model fit result for the white PLA tests, where unfilled squares indicate the mean feed slippage and dashed the fitted model, for speeds 1.5 mm/s, 2.0 mm/s and 2.5 mm/s.

The constants found in the fitting process are given in Table 3.3.

Current Slicer software uses a steady state feed correction factor. The results here can therefore be used to avoid feed slippage or increase the correction factor to reduce the amount of under extrusion. It is also possible, in principle, to use the empirical model to generate extrusion commands for different steady state extruder temperatures and feed speeds.

Table 3.3: Fitted constants for the temperature and speed feed slippage model.

Parameter	Value	Unit
B	186.238	$^{\circ}$C
m_a	2.919	$^{\circ}\text{C} \cdot \text{s/m}$
c_a	-2.578	$^{\circ}$C

Feed Slippage Time Dependency

The feed slippage, however, also depends on time. This effect is visualised in Fig. 3.14, where the feed slippage percentage is drawn against time, for a 2 mm/s extrusion speed at different temperatures.

The time based feed slippage figure shows that the feed slippage is not constant, especially for the first 5 s of extrusion. Failure to extrude any material is seen for the white PLA filament at 190 °C. After around 15 s the feed mechanism reaches a semi-steady state for the 2 mm/s feed speed.

Here are at least two important things to take note of. Firstly, a steady state slicer model for feed slippage cannot correct for the time dependent error. A time based model will have to be implemented in the firmware, since the G-code only specifies the end coordinates (see Section 2.5). The firmware will have to dynamically adjust the print and feed speeds to reduce the volumetric flow error (see for example [139]).

Secondly, a 15 s long extrusion at 2 mm/s, with 2.85 mm filament, will feed 30 mm of filament and create a single track which has a length of 975 mm. It is therefore not practical to wait for the liquefier and feed mechanism to reach this steady state situation.

Feed slippage and the resulting under extrusion can therefore be reduced by either static correction or by dynamic closed loop control. A practical static solution is to simply select process parameters where the slippage is below an acceptable threshold. For example, it can be recommended for the white PLA used here to keep the temperature equal to or above 215 °C and the feed speed below 1.5 mm/s. These are also reasonable limits for normal printing. Note that increasing the temperature is limited by the material degradation limit, post extrusion settling effects as well as the hot end cooling efficiency.

Closed loop control is, however, required if it is desired to increase the printing speed, printer output robustness or accuracy, as discussed in Section 2.9.

3.1.10 Feed Slippage Real-Time Control

The application of feed speed monitoring as a closed loop controller is presented next. The approach used here is to print a test object using a

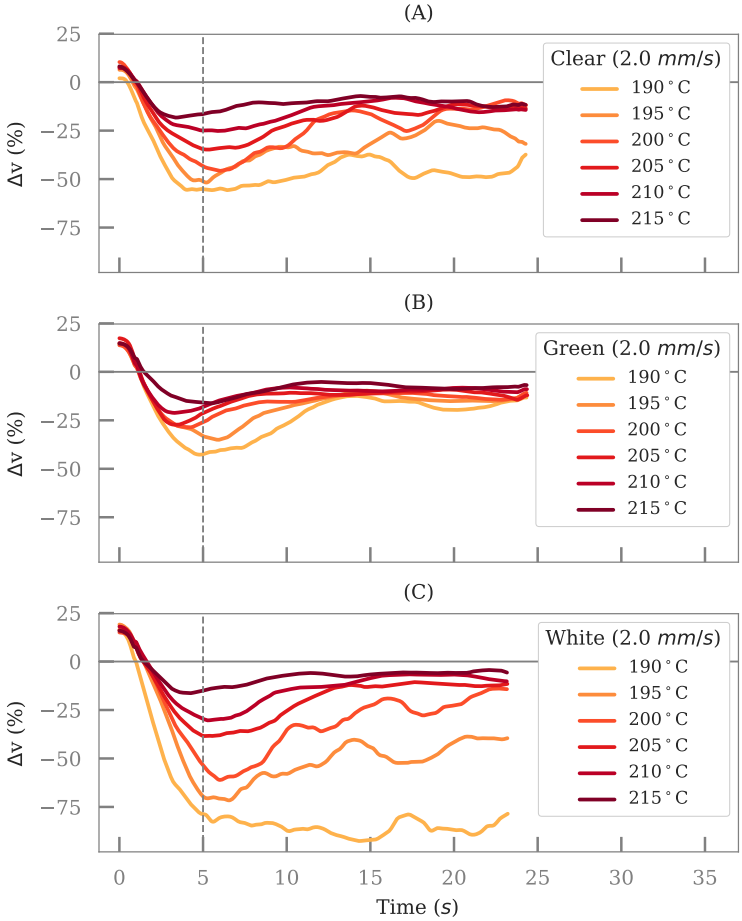


Figure 3.14: Feed slippage versus time at 2 mm/s extrusion speed for (A) clear, (B) green and (C) white coloured PLA filaments. The dashed vertical line indicates the 5 s time point [145].

basic proportional controller.

Controller Interface

The feed slippage is detected by the machine vision application, which communicates with the experiment controlling program. This program determines the control signal, which is sent to the printer through the print server interface.

A flow and speed rate multiplier can be set through the server interface. This is a percentage value, which is multiplied with the G-code command, for the extruder and movement speed in the firmware. The measured slippage is used to determine the value of this multiply command, creating the control signal.

This method is shown in the block diagram in Fig. 3.15. The unmodified printer runs open loop, in which the server reads the object file containing the G-codes, which it sends with a communication protocol to the printer. The printer interprets the G-code command and sets the extrusion motor speed. The loop is closed by sending the speed multiply percentage.

The multiply change command is, however, not seen as a queue command and is only executed after the current movement commands are completed (see Section 2.6.5) and it is only applied during the motion planning and not during the execution stage. This method is therefore only a *test of concept* and needs to be further developed, through a direct firmware interface, to improve the response time. Nevertheless, it was tested on a tailor made object, to gauge if it is feasible to control the feed mechanism with the optical flow feed slippage measurement.

The main program runs three threads, namely the machine vision, master and printer client thread, where the machine vision thread is actually run on a separate processor.

The measured slippage, in percentage, is averaged over time and used directly as the flow and feed rate multiplier. This reduces the slippage by decreasing the extruder speed.

Control Results

A test object was printed twice to test the feed slippage controller. First with the control off and then with it turned on (see Fig. 3.16). The object is a 40 sided polygon with a diameter of 68 mm and wall thickness such that the slicer will build the object perimeter with one track. The object height is 10 mm, with 33, 0.3 mm layers.

The G-code file was modified to create slippage, by printing the first 11 layers at 230 °C, the next 11 layers at 190 °C and the remaining layers again at 230 °C. Large blobs on the wall of the prints are the positions where the extruder waited until the temperature reached a new set point.

The print with the control turned on is seen to be qualitatively much better, with less under extrusion, than the print with the control turned off. The reason for this can be seen by considering the measured data.

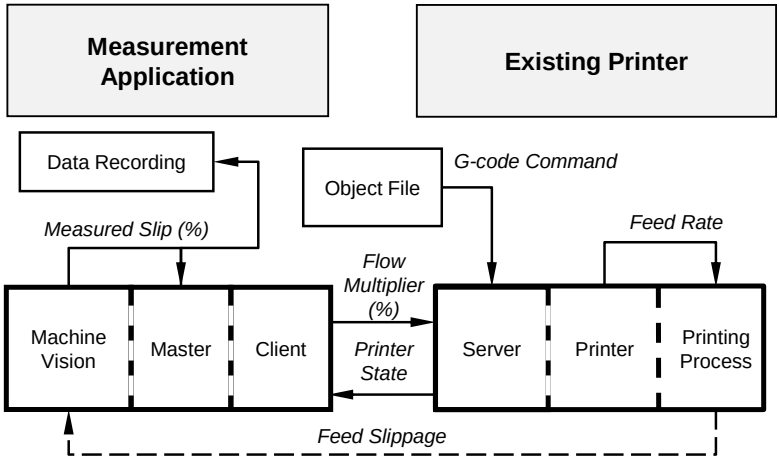


Figure 3.15: Block diagram of the machine vision feed slippage controller [145].

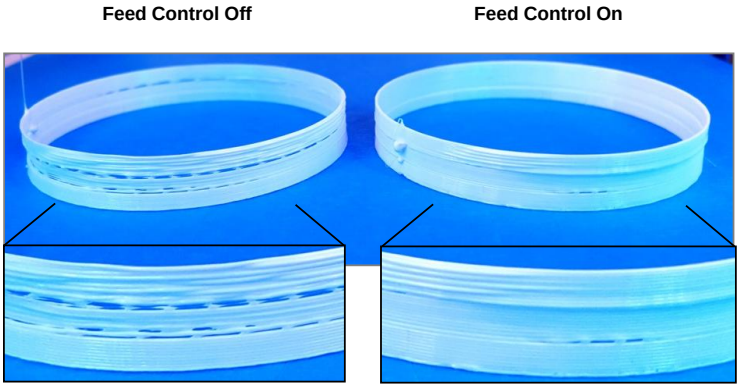


Figure 3.16: Photo of the two feed control tests, with (left) control off and (right) control on. The diameters of the objects are 68 mm.

The temperature versus time profile for both tests are shown in Fig. 3.17, as well as the recorded feed slippage percentage. Feed slippage is only shown for periods of gear movement and additionally filtered to increase readability. Note that the printer was commanded to wait until the extruder reached the set point temperature, before continuing the print. This creates the time gaps in the profiles.

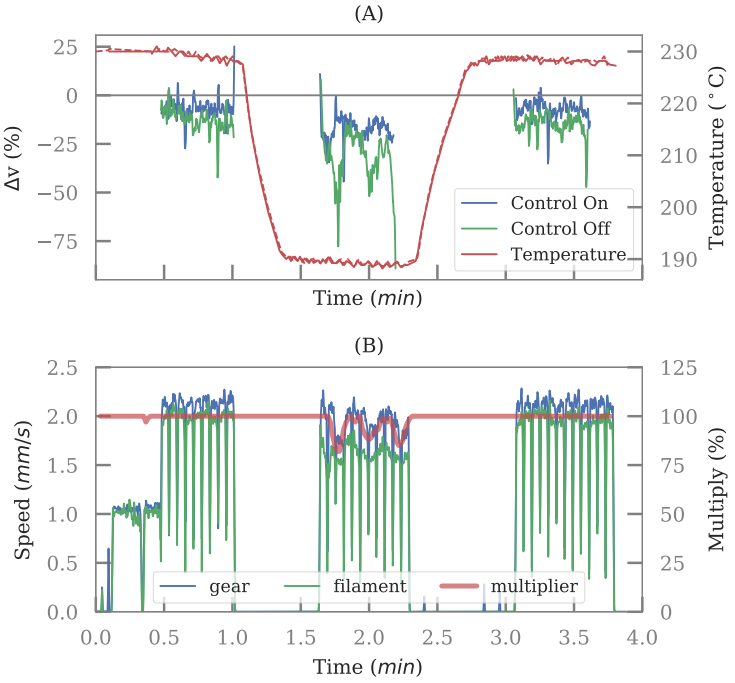


Figure 3.17: Feed control test results, with (A) comparing the feed slippage and temperature for both tests and (B) the speed reduction with the multiplier value, gear and filament speed, during the control test print [145].

Slippage increases, as expected, in the middle segment of the three layer sets and the control loop reacts by reducing the gear speed, in order to regain traction.

The multiplier value during the active control print test is shown in Fig. 3.17 (B), along with the gear and filament speed (the sharp vertical spikes are filament retractions between layers).

The various object sections can be seen in the speed profile. The printer first prints an extruder test line and then the first layer at a

slower speed of 1 mm/s, while the remaining lines are printed with a target speed of 2 mm/s. Furthermore, the gear speed clearly follows the multiplier percentage value, especially in the middle section at the lower temperature.

This test was repeated twice and four objects (two with control on, and two with control off) were weighed. The mass difference with the control on was (30 ± 1) mg, while the mass difference between the two objects printed without active control was (21 ± 1) mg. This shows the tests have reasonable repeatability between prints, given the calculated object mass of 1903 mg.

The average difference between the prints with and without active control is equal to (198 ± 1) mg, which is a 14 % mass increase by turning the control on. Active control therefore significantly reduces under extrusion. The test further confirms that such a controller is feasible and can improve the final print result.

3.2 Retraction Improvement

Precise volumetric flow rate is critical to achieve high quality parts. This includes the start, steady state and end-sections of the extrusion. A method to measure the input flow rate is described in Section 3.1. The exit flow rate is, however, harder to measure, but an important quantity during extrusion start and stop, when the flow rate is not a steady state level (see Section 2.8.10).

A single extrusion can be divided into three dynamic extrusion segments the: (1) extrusion start, (2) feed speed change and (3) retraction. The purpose of retraction is to reduce the pressure in the liquefier as fast as possible, in order to stop the unwanted flow of molten material or *oozing*.

Retraction is studied in this section with a cost effective experimental method, which can measure the exit flow rate during test extrusions as well as the filament feed slippage. The RF1000 printer bed level sensor is also redeployed as a liquefier pressure sensor. These measurement tools are used to study different retraction techniques and presents an objective way to compare different printer settings without printing a large set of test objects.

This section is based on the work presented by Greeff and Schilling [180]. The exit flow measurement and pressure sense methods are presented first, which includes an uncertainty estimation for each method. The off-line retraction experimental plan and results are discussed next, followed by a test object experiment, which is used to compare the in-process measurement results with post-process, qualitative, measurements.

3.2.1 Exit Flow Measurement Method

The gear, filament and exit speed, as well as the entry and exit filament width, are measured optically with two different USB microscopes, using the image processing presented in Section 3.1.

The maximum speed, which can be measured, is limited by the frame rate of the cameras. The feed mechanism camera achieves 30 FPS, while the exit flow camera could only achieve 26 FPS. In principle, the cameras could also be swapped around, so that the slightly faster camera measures the exit speed. This, however, was not necessary, as the frame rate was fast enough for realistic printing speeds, with this printer.

The material exiting the nozzle is moving approximately 32 times faster than the input speed, with a filament diameter 2.85 mm and a nozzle diameter of 0.5 mm. The exit speed field of view should therefore be large enough to detect the faster exit speed. Furthermore, the extruder camera can track the gear tooth markings on the filament or the filament texture itself, but this is not possible for the exit flow camera. The extruded material is not only thinner and faster, but also smoother.

The solution used here is to generate traceable marks on the extruded material in real time. Other methods, such as marking the filament before extrusion or creating a speckle pattern with a laser pointer beam reflection were also tested. Pre-marking the filament did not work since the colour was smeared over a long length of extruded material, due to the 32 times speed-up factor. This, therefore, only created a gradual change in the colour of the exit flow, which could not be used to determine the flow rate.

Forcing small salt crystals into the filament, before extrusion, was also tested, aiming to create a more textured output flow, which could be detected if a laser beam was reflected of the surface. This also did not give reliable results. Modifying the filament before extrusion is also not ideal, since this could change the flow characteristics.

A patented velocimetry method was mentioned in Section 2.9 and the only known real-time exit flow rate measurement technique. This, however, requires a special liquefier and expensive instrumentation. Most studies, therefore, extrude a length of filament and then either weigh it or measure the track width, to determine the flow rate, but this is not an in-process (real-time) measurement.

In this section the aim is to measure the real time flow rate of test extrusions, with a cost efficient and open source solution. Test extrusions imply free air extrusion, i.e. it cannot be used during actual object printing. The concept is to generate points on the exit flow, which can be optically tracked, using the optical flow algorithm presented previously.

These marks are generated with ink jetting, using an obsolete, but still available, 96 DPI (dots per inch) inkjet cartridge and controlling hardware from the open source project, which is called *InkShield* by Lewis [181]. The implementation is shown in Fig. 3.18.

The inkjet rate must also be controlled to accommodate for changing speeds, since the jetting rate will determine the spatial period of the dots, relative to the exit speed. The rate is adjusted by the main controller application, which based on the measured feed gear speed. This ensures that there are always dots spaced approximately equally over the exit flow ROI.

The diameter of the extruded material is also estimated using the same method for the filament width. This is combined with the exit speed to

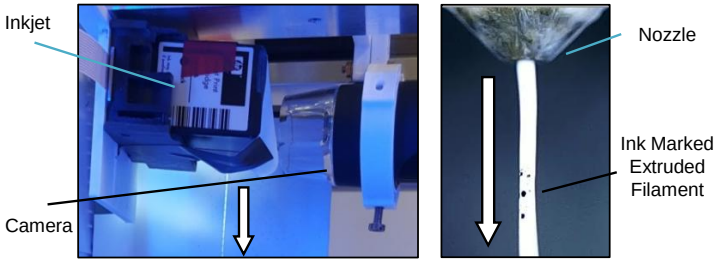


Figure 3.18: Exit flow measurement instrumentation with the inkjet cartridge and USB microscope camera (on the left) and a photo taken by the camera of the ink marked extruded filament (on the right). Large arrows indicate the flow direction.

determine the volumetric exit flow rate.

3.2.2 Exit Flow Measurement Limitations

This method only works for test extrusions, as noted before, since there is no space for the exit flow measurement instruments during normal printing.

The effective removal or conveying of the extruded material, without affecting the flow or the position of the material in the image ROI, is by far the greatest issue with this method. Furthermore, the weight of the extruded filament also pulls the hot material down, affecting the speed and width measurements. Note that in this work the weight effect is not considered.

The print bed was positioned at the maximum distance from the nozzle, allowing the filament to flow downwards and then, ideally, naturally, form a cylinder on the bed. This was not always the case. For example, the extruded filament sometimes curled up, back onto the nozzle, directly after extrusion.

Another issue was the formation of *pillars* instead of cylinders. This is when the solidified material reaches the bed and instead of gently bending down onto the surface and building a cylinder, it forms a rigid rod, which at some point falls to the side, pulling the flow out of the image ROI.

Nevertheless, measurements and results could still be obtained using this cost effective technique. The same method and parameters used to estimate the uncertainty for a 2 mm/s feed speed was applied to the exit volumetric flow rate ($\dot{Q}_e$) and the extruded volume (V_e) (see Section 3.1.8). The uncertainty for $\dot{Q}_e$ is estimated as $\pm 0.51\,\mathrm{mm}^3/\mathrm{s}$ and for V_e it is estimated as $\pm 1.3\,\mathrm{mm}^3$, with a coverage factor of two (95 % (normal)).

3.2.3 Liquefier Pressure Drop Measurement Method

The RF1000 liquefier is mounted on a plate, which is supported by two cantilevers. Each cantilever is fitted with strain gauges and fixed to the printer X-axis carriage. The feed mechanism is also mounted to the same carriage. This allows for the liquefier to slightly bend the cantilever beams up or down, depending on the resulting force on the liquefier Fig. 3.19.

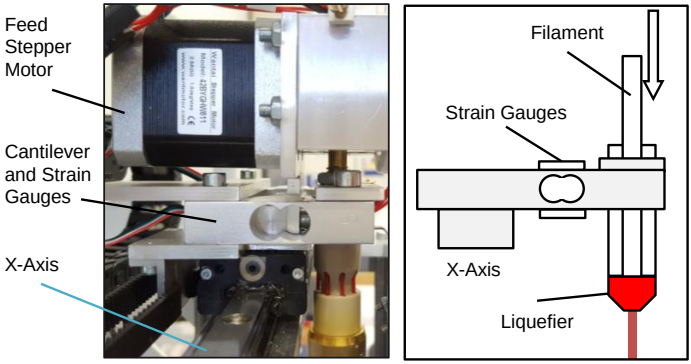


Figure 3.19: Liquefier pressure measurement with the printer bed level force sensor. A photo (left) from the side of the feed stepper motor and X-axis shows how the liquefier is mounted on the cantilever. The diagram (right) shows how the force exerted by the filament on the melt translates to the bending of the cantilever, which is measured by the strain gauge bridge.

The strain gauge bridge converts the cantilever bending to a change of resistance, which is measured by the RF1000 electronics using an analogue to digital converter (ADC). The RF1000 printer uses this as a bed level sensor, where the bed bends the cantilevers upwards, as the nozzle makes contact with the surface.

In this work the printer firmware was modified to re-purpose the strain gauges to measure the liquefier pressure. A function was inserted, which reports the strain values at a set interval of 10 Hz, similar to the way the printer temperatures are reported (see Section 2.6.5). Unfortunately, faster rates were not possible without affecting the print quality.

The strain gauge reading (x_{counts}) was verified by loading calibrated weights on a rod, which was mounted in the place of the liquefier. The weights were used to generate a mass from 0.1 – 4 kg, in steps of 0.1 kg. Each mass piece was applied for 5 s.

The maximum range (maximum minus minimum), over all the measurements, was 5 counts (approximately 2.9 g), while the average range was 2.4 counts.

A linear fit according to Eq. (3.9) was applied to the data, which converts the ADC counts to mass $M\,(x_{counts})$, with the two fit coefficients m_{fit} and c_{fit}. The gradient was found to be 0.584 g/count with c_{fit} equal to 0.0 g.

$$M\,(x_{counts}) = m_{fit}x_{counts} + c_{fit} \tag{3.9}$$

The mass is converted to pressure (ΔP) using Eq. (3.10), with $F = ma$, R_f the filament radius and the gravitational acceleration ga.

$$\Delta P = ga\,\frac{M\,(x_{counts})}{\pi R_f^2} \tag{3.10}$$

3.2.4 Liquefier Pressure Drop Uncertainty

The same method and parameters used to estimate the uncertainty for a 2 mm/s feed speed was applied to the pressure drop measurement (see Section 3.1.8) with Eq. (3.10). Assumptions include that the filament is perfectly round, the melt is incompressible and that the strain gauge and cantilever beam dynamic effects can be neglected.

Table 3.4: Pressure drop uncertainty parameters for a 2 mm/s feed speed, with abbreviations: Exp. (Expanded), Unc. (Uncertainty), Dist. (Distribution), Rect. (Rectangular) and Cov. Factor (Coverage Factor).

Factor	Distr.	Value	Unc.	Unit	Note
m_{fit}	Normal	0.584	0.004	N/count	Exp. Unc. Cov. Factor: 1
x_{counts}	Rect.	5570	2.5	pxl	Half width limits
D_f	Normal	2.85	0.070	mm	Exp. Unc. Cov. Factor: 2

The uncertainty for ΔP, with the required additional parameters given in Table 3.4, is estimated as ±0.26 MPa or 13 %, with a coverage factor of two (95 % (normal)).

3.2.5 Retraction Experimental Plan

White PLA, with a nominal diameter of 2.85 mm, was used for all the retraction tests, with an extrusion temperature of 215 °C. Different retraction methods were applied and the results recorded to investigate the retraction process.

Before each extrusion experiment the *nozzle prep routine* was executed, which is the same set of commands that the RF1000 normally performs, before starting an actual print. It involves three extrusions: (a) fast nozzle infill (13 mm at 13.33 mm/s), (b) longer liquefier preparation extrusion (25 mm at 0.52 mm/s) and (c) finally a retraction (1 mm at −16 mm/s).

A set of retraction tests was executed after completing the nozzle prep routine. This set included different retraction techniques. After each retraction, a three step post retraction routine was executed. This involves (a) switching off the extruder stepper motor after a time delay and extruding filament again after another short delay. This is done by (b) first extruding a 1 mm liquefier *refill* at 16 mm/s and (c) then continuing the extrusion for another 13 mm at 0.75 mm/s.

The test files were generated using a python script, which created a list of G-code commands to perform the extrusion and retraction moves. A single test was performed by sending the whole list to the printer through the print server interface. The measured data were combined by interpolation to generate a single data file, containing the temperatures, pressure, speeds and widths, per time point.

The different retraction methods tested here are summarised in Table 3.5. The values for the *Normal Retraction* are derived from standard RF1000 printer settings, as supplied by the manufacturer at the time of purchase.

Table 3.5: Retraction test plan

Retraction Name	**Length of Filament** mm	**Speed** mm/s	**Legend Key**	**Dwell Time** s
No Retraction			None	0.00
Normal Retraction	1.0	-16.0	V16	0.06
Speed Variation	1.0	-8.0	V08	0.13
		-12.0	V12	0.08
		-20.0	V20	0.05
Length Variation	0.5	-16.0	L05	0.03
	1.5		L15	0.09
Multi-Step	*Repeated 3 times*		Multi-Step	0.39
	1.0	-16.0		

Analysis of the retraction data required the segmentation and labelling of each test data file. A single test run executed the preparation routine

as well as a set of extrusion and retraction tests. Each retraction was performed with different parameters, as defined in Table 3.5.

For a *length variation* test this means that at least two retractions were required. The data was analysed to determine the time segment of each retraction experiment. Each segment was labelled, time shifted (to align the datasets due to small differences) and averaged with the other tests, to determine a representative profile for each retraction type.

The multi-step method is a non-standard retraction. This was included to investigate the pressure reduction limit. It consists of a repetition of three *normal* retractions, each followed with a short refill extrusion of 0.1 mm at 0.5 mm/s.

The concept of the multi-step method is that the feed mechanism can regain traction, since feed slippage occurs during fast retractions. This is, in general, not a problem since the retraction length is relatively short. It does, however, limit the actual length which is retracted. Furthermore, it is thought that the retracted semi-molten filament has more time to solidify with the multi-step method and can therefore be pulled back higher up in the liquefier, increasing the total pressure reduction.

3.2.6 Result of the Retraction Experiments

The measured data for a single *length variation* test is shown in Fig. 3.20. The liquefier preparation routine takes about 55 s to complete, after which four test extrusions can be seen. The use of the preparation routine is motivated by the data, as it takes a relatively long time before the flow rate and the pressure reaches a semi-steady state situation.

A dashed line shows the G-code commanded flow rate and the negative vertical spikes indicate retractions. Note that negative volumetric flow is not possible here. The volumetric input flow follows the commanded flow, with some under extrusion, which is also affected by the filament diameter.

As expected, the liquefier pressure also follows the input flow trajectory. The pressure decrease and the material flow, during and after a retraction, depend on the applied retraction technique (see Fig. 3.20 (B)).

At certain points it can be seen that the exit flow rate significantly deviates from the input flow rate. These are locations where the extruded material removal caused speed detection errors, as discussed in Section 3.2.2. Several measurement sets were therefore taken, aligned and averaged to compensate for these effects.

Material *ooze* flow is considered first, where ooze is the unwanted flow of material after stopping the extrusion. The average ooze flow is shown in Fig. 3.21 with the flow uncertainty estimation determined in Section 3.2.2. No retraction (*None*) has the slowest decrease in the ooze flow rate, while the L15, *length variation* test, achieved the fastest decrease rate.

The non-zero, steady state, ooze flow can be attributed to the weight of the extruded material dragging itself downwards. Both the L05 and multi-step methods achieve the slowest ooze flow at the minimum point ($0.4\,\mathrm{mm}^3/\mathrm{s}$), while stopping the extrusion without retraction has a mini-

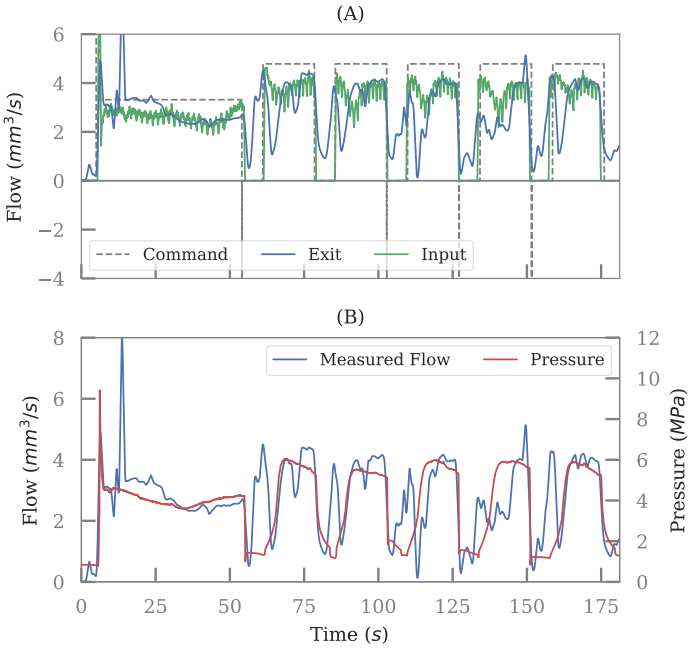


Figure 3.20: Results measured for a *length variation* test, with (A) comparing the volumetric input, commanded and exit flow and (B) the exit flow and measured liquefier pressure.

mum ooze flow of approximately 1 mm^3/s. All the applied retractions do, however, reduce the ooze flow, indicating that retraction is beneficial to the printing process.

Furthermore, Fig. 3.21 not only indicates that increasing the retraction speed and decreasing the retraction length, reduces the ooze flow rate, but also that all the methods have an approximately equal decreasing flow rate slope, except for the multi-step method. This slope can, however, be influenced by the averaging and data filtering. Note that the multi-step method involves a short refill extrusion after each retraction step. Extra ooze is therefore generated when the refill length is too long.

The volume of the oozed material is calculated by integrating between the dashed vertical line and the ooze end point, for each test and the result is shown in Fig. 3.22, with the uncertainty estimate derived in Section 3.2.2.

The uncertainty bars limit detail conclusions, but the L15 *length variation* technique yields the smallest oozed volume, while applying no retraction results in the largest amount of unwanted material flow.

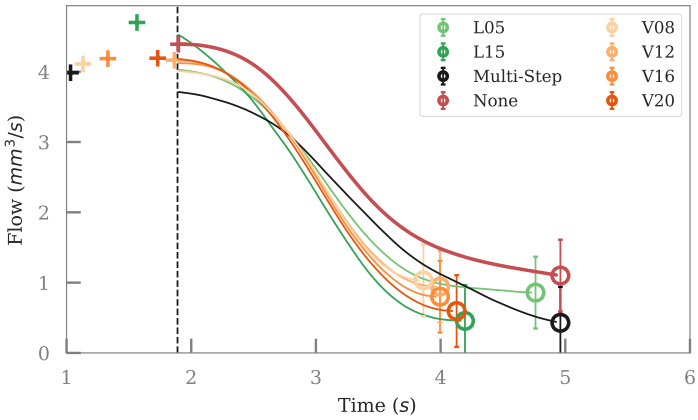


Figure 3.21: Material flow after extrusion stop. Pluses indicate the start of each profile (maximum point), circles the end (minimum point) and the dashed vertical is the line mean starting time point.

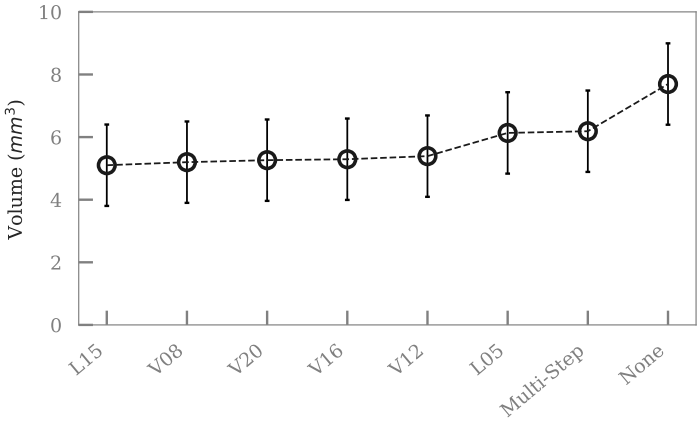


Figure 3.22: Calculated volume of oozed material after extrusion stop.

The retraction test results are investigated further by considering the pressure drop, since the main motivation for using retraction is the assumption that the melt pressure after extrusion completion is the biggest contributor towards unwanted material flow. The average pressure drop against time for the different retraction tests are shown in Fig. 3.23.

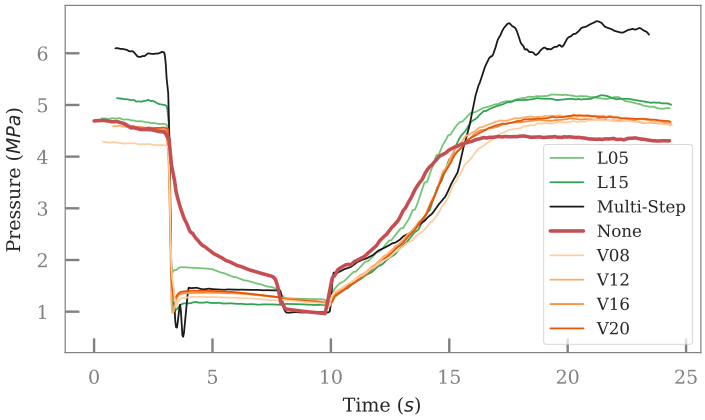


Figure 3.23: Liquefier pressure versus time during the retraction tests.

The first sharp drop in pressure indicates the retraction event at around 3 s. A second pressure drop is seen at 8 s. This is when the stepper motor is disabled. Extrusion is then restarted at 10 s, where all the curves show a step increase, after which different step responses can be distinguished.

After each retraction completion a small and fast pressure increase is seen. One explanation for this is that a returning force acts on the system due to the cantilever stiffness and the weight of the extruder, similar to the step response of a compressed mass-spring system.

An unexpected force decrease, after disabling the stepper motor, is seen in the *none*, L05 and multi-step retraction tests, which indicates that the motor was still applying a compressive force on the melt (due to the holding torque). This force decrease can therefore be used as an indication of how successfully the filament was detached from the melt. (Note that disabling of the stepper motors during printing is not allowed, since the printer will loose the home position.)

The remaining force for the multi-step method can be explained since it performs a refill move after the last retraction, while the L05 retraction length is probably too short. The nozzle tip length was found to be 1 mm in Section 2.8.3, which motivates the use of retraction length equal to or longer than 1 mm.

The detailed pressure response for the different retraction techniques are shown in Fig. 3.24. The multi-step technique is seen to reduce the

pressure in three steps and achieves the minimum pressure, as indicated by the unfilled circle.

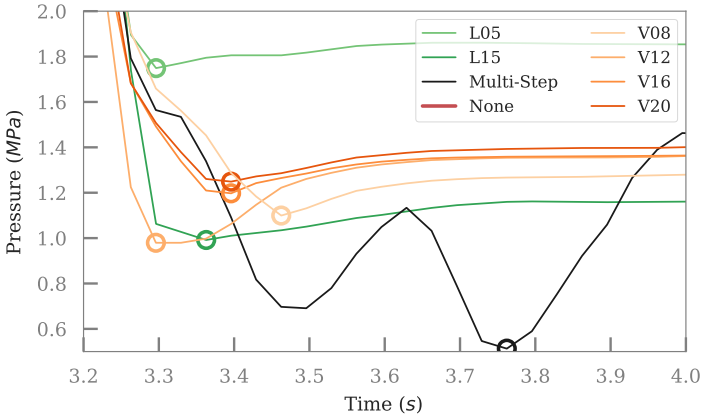


Figure 3.24: Detail of the pressure response for each test, with unfilled circles showing the minimum point.

The average initial pressure of each test before the retraction test was, however, not the same. This is also seen with the extrusion response after the retraction and delay time in Fig. 3.23. No retraction has the lowest re-extrusion pressure, followed by the velocity and length variation tests. The multi-step technique has by far the highest re-extrusion pressure.

It is still unclear why this is the case. Furthermore, the same re-extrusion command was used for all the tests. In practise, the amount of refill will depend on the retracted length.

Another possible explanation is the temperature of the filament. The filament will have the highest temperature if no retraction is applied, which will require the least amount of pressure to extrude. The multi-step method, with improved traction and the largest total retraction length, should have the coldest filament temperature, requiring significantly more pressure to extrude. This does not however explain the apparent difference between the length and speed variation methods.

The pressure drop percentage was also calculated to compare the different methods, by assuming an identical initial pressure with $(\Delta P_{max} - \Delta P_i)/\Delta P_{max}$, where ΔP_{max} is minimum pressure achieved, without applying any retraction and Δ_i is minimum pressure for each method.

The results are shown in Fig. 3.25, with the uncertainty estimate determined in Section 3.2.4.

This quantifies that the multi-step method achieves the biggest pressure decrease. The results of the other tests are also interesting. Increasing the retraction speed, does not necessarily increase pressure drop, contrary

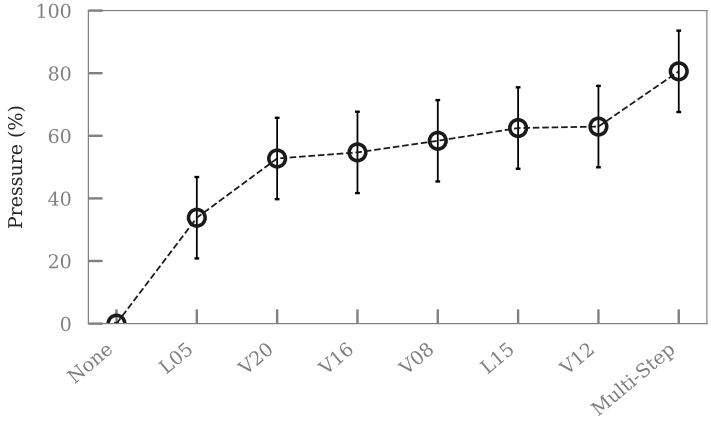


Figure 3.25: Pressure decrease percentage, as compared to minimum pressure when no retraction was applied.

to the expectation, since the feed mechanism slippage will also increase, reducing the actual amount of filament reversed. Increasing the retraction length, however, does increase the pressure drop percentage.

3.2.7 Verification with a Test Object

The in-process retraction test results were verified by applying the methods to a test object and optically measuring the resulting material ooze.

A base object file was created using CAD and sliced. The sliced file was then modified using a Python script, which added G-codes so that the different retraction methods could be tested with a single test print (see Fig. 3.26). The concept here is to extrude a thin wall, perform a specific retraction test and then move away to print the next thin wall. Each retraction used two thin walls, as shown in the figure. The length of the wall is important since it ensures that the feed mechanism and liquefier pressure achieves a semi-state condition.

A non-print move, after a retraction will, however, also create stringing, which is the formation of thin strings pulled out of the molten track or ooze material. The amount of stringing should increase with an increase of oozed material, but also with decreasing dwell time (see Table 3.5), which is the time required to complete the retraction. An increased dwell time will allow the track material to solidify more, reducing the stringing. The nozzle design will, however, also influence the stringing, see Section 2.8.6.

Photos were taken with a USB microscope camera at each of the retraction points of the test object and are shown in Fig. 3.27.

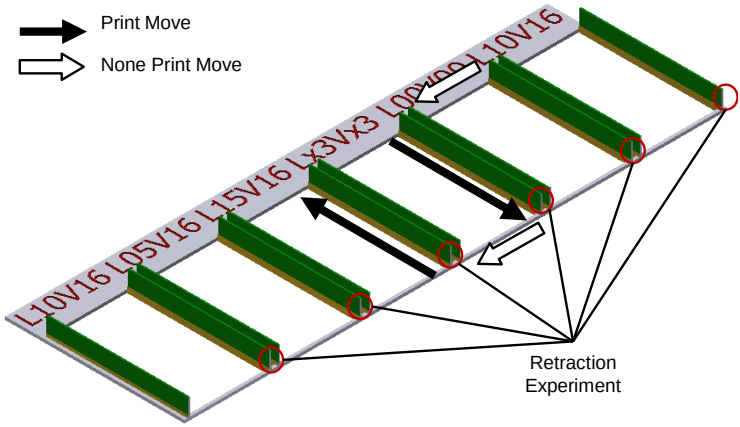


Figure 3.26: Retraction test object design. The base part (grey) was designed in CAD, with embossed letters (red) indicating the experiment code. The green walls were generated by the Python script. Filled arrows are print moves (extrusions), unfilled arrows are non-print moves and red circles are retraction experiment locations.

The multi-step method achieves qualitatively the best results of all the methods, while performing no retraction results in most stringing (front view) and the largest track bending radius (top view). Again, it is seen that slower retraction speeds give better results (compare for example V08 with V20). Apart from the feed slippage, the main reason for this is the longer dwell time (see Table 3.5).

The stringing length of each retraction test in Fig. 3.27 was assigned a value from zero to five, where zero is no stringing (multi-step method) and five is the maximum (no retraction). These stringing values are graphed against the dwell-time in Fig. 3.28 (A) and confirms that the qualitative stringing length decreases with increasing dwell time. It is also compared with the maximum pressure decrease percentage in Fig. 3.28 (B). This shows that the stringing length is relatively independent of the pressure drop.

The manufacturer supplied standard retraction setting (V16, see Table 3.5) is shown with a vertical dashed line. It is approximately in the centre of the spread of values.

3.2.8 Discussion of Retraction Results

The cost effective method to measure the input flow rate presented in Section 3.1 was further developed here to measure the exit flow rate for test extrusions, with in-process inkjet marking of the material flowing out

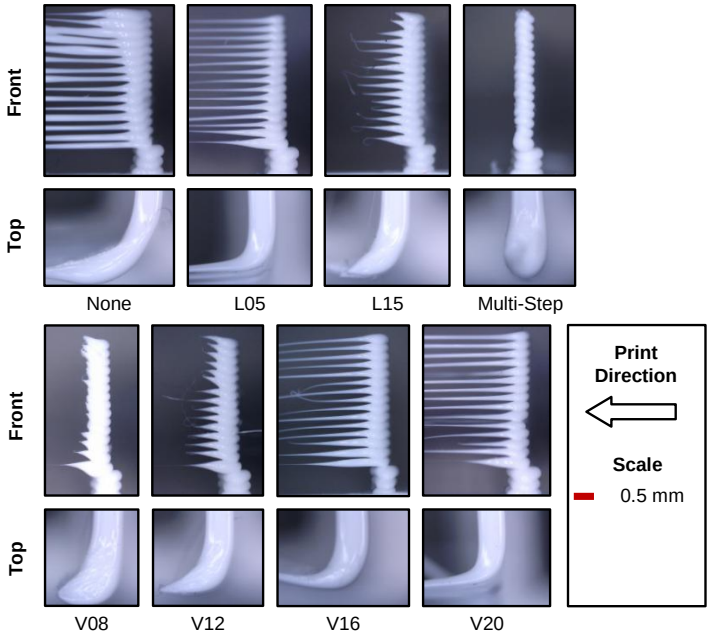


Figure 3.27: Photos of the retraction test object stringing length, at the retraction locations, taken from the front and from above.

of the extruder. The jetting rate is adjusted in real time, which enables the measurement of both fast extrusion flow and slower, post extrusion or ooze flow. A method to measure the pressure drop over the liquefier was also presented.

The presented method can be used to verify if an extruder is working properly, compare different extruder designs or to determine the optimal configuration parameters for a new feedstock.

The ooze flow, or unwanted flow after extrusion stop, was investigated in this section, by applying different retraction techniques. This included retraction with standard parameters, no retraction, as well as a special multi-step method. The retraction techniques were investigated using an automated test bench and qualitatively compared with a script generated test piece.

This section finds that the dwell-time at the retraction location is a significant contributor towards the stringing length. This length is, however, relatively independent of the pressure decrease, which was not expected.

A minimum retraction length of 1 mm can also be recommended, which

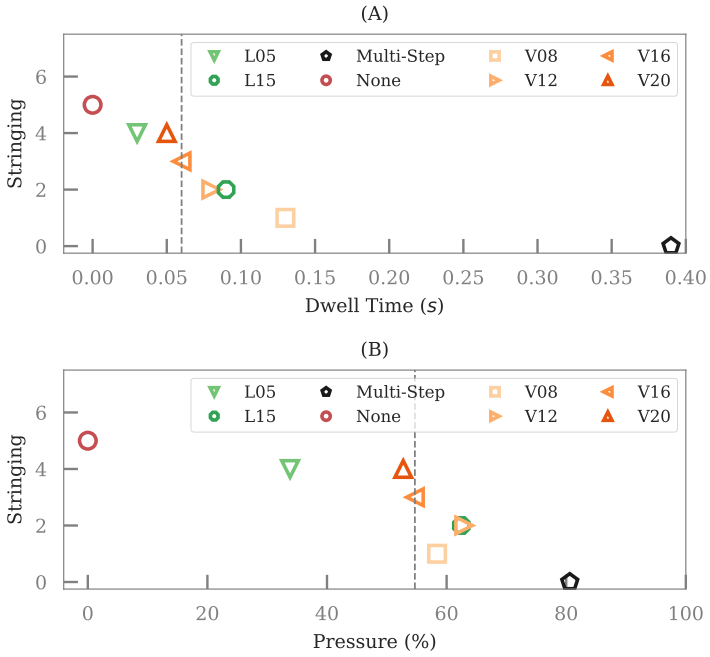


Figure 3.28: Comparison between (A) the qualitative stringing length and the dwell-time and (B) the stringing length and pressure decrease percentage. Vertical dashed lines indicate the standard retraction (V16).

is equal to the nozzle tip length. Interesting force dynamics were also measured during the retractions and extrusions. This was briefly discussed in this section and will be considered in more detail in Section 3.4.

3.3 Single Print Optimisation

The main contribution of this section is a practical, easy-to-use and universal method, called Single Print Optimisation (SPO), for executing a Design of Experiments (DOE) run table (or even more than one table) with a single print run. SPO significantly reduces experimentation time through semi-sequential object printing, with super-layer merging.

The review of FFF optimisation (see Section 2.10) found that there are many process parameters, which affect the final part geometric accuracy, mechanical properties and cost. It was also found that these parameters can be efficiently optimised using DOE.

The proposed SPO method is practically applied in this section to

optimise the dimensional features of a specially designed test part by using Design of Experiments (DOE). A bottom-up approach is used by considering the execution accuracy of the **G1** print command. The section contributes further by presenting a test part, which combines both Design-for-Metrology and Design-for-Additive-Manufacturing (DfAM). This test part is designed in such a way that it does not require expensive 3D measuring instruments.

Note that Design of Experiments (DOE) is not new to Additive Manufacturing or in Material Extrusion processes, such as FFF. The novelty of this work is the practical realisation of Single Print Optimisation (SPO), without the need for expensive measurement technology or complex optimisation algorithms.

The SPO method presented here can be applied by any user, researcher or Original Equipment Manufacturer (OEM), who wants to improve the FFF process, output or quality assurance. This is also the motivation for the inclusion of the test part design. Personal experience has found that certain FFF printer manufacturers and users use a very expensive trial-and-error method for parameter optimisation, by using large (long print time) qualitative test pieces. Furthermore, these test objects are also not designed to be quantitatively measured with accessible and easy to use measurement devices.

This section is based on [182], which was developed from the following goals: (1) the optimisation of slicer parameters by comparing the commanded G-code extrusion with the actual print result, (2) validation of the SPO method and (3) the design of a test piece, which (a) enables the first goal, (b) does not require a non-standard slicing process and (c) is measurable with easy to use, accessible and affordable instruments.

3.3.1 The RF1000 Geometry and Slicer

The RF1000 printable XY area is 245 mm by 250 mm and the Z height is 200 mm. The bed is moved in the YZ-plane and the nozzle, which has a diameter of 0.5 mm, is moved in the X-axis. A clearance of approximately 7 mm is created by the nozzle protruding below the X-axis frame. This allows for sequential printing of 7 mm tall objects, without the risk of nozzle collision in the Z-direction. Sequential objects taller than this must be merged into *super layers*.

The printer uses firmware based on the Repetier FFF firmware branch and the *Slic3r* (version 1.2.9) slicer [42] is used throughout. Note that the *auto cool* option in the slicer was turned off, to ensure that the print speed is not automatically reduced by the slicer to increase layer cooling time.

3.3.2 SPO Introduction

An opportunity to combine DOE with single run sequential printing was found in Section 2.10. This introduces the idea of Single Print Optimisation (SPO), where the goal is to rapidly and systematically identify significant process factors and optimise them in a single print run.

This overcomes the limitation of concurrent object printing, since it allows for interlayer effect testing. The printing time cost of using several print runs is also overcome, as the pre- and post-processing steps for an object needs to be performed only once.

The SPO process is shown Fig. 3.29, with number and arrows to indicate the print sequence.

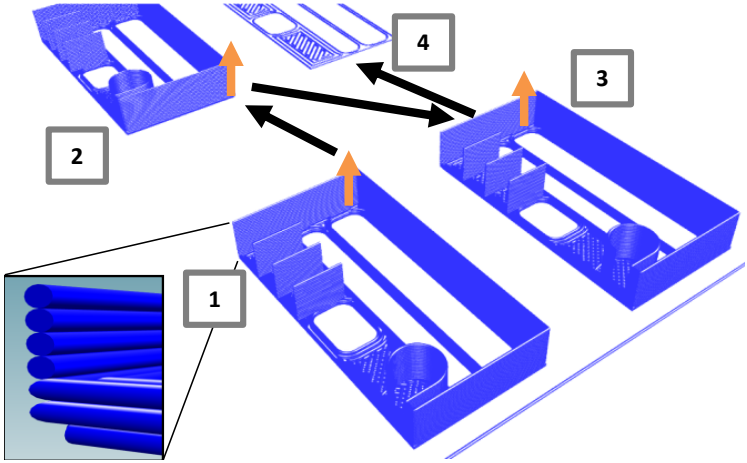


Figure 3.29: The print sequence visualisation of Single Print Optimisation (SPO), where each object is sliced with a different parameter set and is printed sequentially, as indicated by the arrows and numbers. The blue structure is created by visualisation of the G-code commands and the inset shows the details of a few print moves.

Single run sequential printing is already possible with the slicer program used in this work. The difference here, however, is the ability to merge differently sliced objects into a single file, where each object is also a combination of *standard layers* and *non-standard layers*. The first n layers, in this case three, are printed with the standard parameters. The remaining layers are printed according to the DOE run table, by adding the non-standard layers.

G-code file merging also creates *super layers*. Each object is printed only to the maximum clearance height, if the object is taller than the Z-axis clearance. This is called a super layer. After completion of a super layer, the printer continues again with the first object, until all the objects are completed.

The main process steps used to create the final, merged, G-code file are summarised in Fig. 3.30. The SPO output file can be printed just like a normal G-code file.

The algorithm ensures that the nozzle does not crash with other objects and that the printer state is initialised correctly when moving to a different

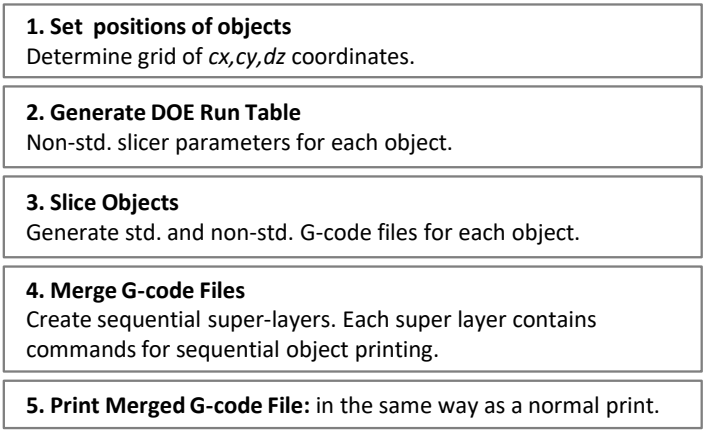


Figure 3.30: Diagram of the Single Print Optimisation process flow.

object. The state, including bed temperature, acceleration settings, etc. should be the same as if object printing was not interrupted. Here is one caveat: most slicers perform a nozzle fill and a preparation extrusion routine, before printing the object. This extrusion is not performed before the printing of each object is started. The standard layers are, however, assumed to be sufficient for bringing the machine into this semi-steady state.

The process of generating the grid objects by defining the slicer parameters, slicing each object and finally merging them again, is coordinated through a GUI (Graphical User Interface). The GUI, scripts to realise each process step, as well as the code to analyse the results are written in the Python programming language. This keeps the work open and accessible. The DOE table is stored in a Microsoft Excel workbook and the measurement results can also be directly typed into this file and imported for result analysis. In principle, it is possible to replace the workbook with any tabular file format.

3.3.3 Design of Experiments Run Table

DOE can be used to derive a model, which describes the relationship between the response variables and the input factors. Here the response variables are the object height, extrusion width and the stringing length. The input factors are the object position, temperature, feed speed and the retraction speed. The object height is measured over the first three layers at the standard settings. This assumes that the object height is only affected by the bed levelness and therefore the part position on the bed. Conversely, the extrusion width and stringing length are considered

to be independent of the part location.

This also assumes constant filament material, condition and diameter, where constant implies a normal deviation from the nominal value. Furthermore, it is assumed that the track height is only determined by the parallelism between the build plane and the nozzle horizontal plane. Pre- and post processing are not considered in this work. The accuracy of selected digital object format, for example the STL file, effectiveness of the slicer, or the effect of post-process solidification are not considered and assumed constant.

Based on this discussion a standard DOE model was selected (see Table 3.6). A *Box-Behnken* quadratic response surface design was selected (see [183] for details), since 18 test objects with sufficient clearance can be fitted inside the build area. The design requires only 15 test objects, with three levels for each factor and three centre runs $[0, 0, 0]$. Other classic designs like CCC, CCI or CCF (Central Composite Circumscribed, Inscribed or Face centred) require at least 20 runs [183].

The run sequence was randomized and three additional runs were added, namely one centre-run and two extreme runs ($[1, 1, 1]$ and $[-1, -1, -1]$). Furthermore, the first two runs are centre-runs, as well the last and middle run.

This run table also allows for the interaction effects and assumptions to be gauged, for example to test whether the retract-speed is the main influence factor of the stringing length.

The factor ranges, given in Table 3.7, were selected based on the manufacture recommended settings for the printer and the filament.

3.3.4 Model Fitting

Ordinary Least Squares (OLS) regression was used to fit the input factor values for the DOE table to the three response measurands. The *Statmodels* Python package was used to perform the regression analysis [184].

A full quadratic model was fitted using Eq. (3.11).The factor with the largest p-value was iteratively removed from the model, until the significance of all the remaining factors were below 0.05 for a 95 % confidence level.

$$\begin{aligned} y = {} & \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \\ & + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 \\ & + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{33} x_3^2 \end{aligned} \tag{3.11}$$

3.3.5 Track Width as a Response Measurand

This work takes the bottom-up approach and treats the FFF end product as a sequence of linear extrusions, or tracks, deposited in the XY plane, as commanded by the *G1* print move (see Section 2.5.2).

Table 3.6: DOE run table with 18 runs and three factors.

Object Coordinates			Coded Factors		
Run	**cx**	**cy**	**x0**	**x1**	**x2**
0	30	45	0	0	0
1	30	123	0	0	0
2	30	201	0	-1	1
3	65	45	1	0	-1
4	65	123	0	1	1
5	65	201	0	-1	-1
6	100	45	1	0	1
7	100	123	1	1	0
8	100	201	-1	-1	-1
9	135	45	0	0	0
10	135	123	0	1	-1
11	135	201	1	-1	0
12	170	45	-1	0	-1
13	170	123	-1	1	0
14	170	201	1	1	1
15	205	45	-1	-1	0
16	205	123	-1	0	1
17	205	201	0	0	0

Table 3.7: Factor ranges

Code	**Parameter**	**-1**	**0**	**1**	**Unit**
x0	Perimeter-Speed	25	40	55	mm/s
x1	Temperature	200	215	230	°C
x2	Retract-Speed	8	12	16	mm/s

The slicer determines the required track width to form the part using a model which assumes that the constant cross section track area is a rectangle with two half circle sides (see Section 2.8.14), where the track area A_t is a function of the layer height h_z and track width W.

The slicer further assumes that the input flow rate is equal to the output flow rate with $L_{xy}A_t = E_f A_f$, where the displacement of the track in the XY plane is L_{xy}, the length of filament to feed is E_f and the cross-section area of the filament is A_f. It then determines the amount of material to be fed into the liquefier to achieve the required width, given a constant filament diameter of D_f. This neglects dynamics, such as start up acceleration or material compression in the liquefier.

The slicer can vary the width of the tracks, i.e. the amount of material fed into liquefier, to achieve different results. For example fine features on the object edge or low resolution object infill can be extruded. The length of filament to feed E_f is determined with Eq. (2.62) ($E_f = 4L_{xy}A_t/(\pi D_f^2)$).

The slicer assumes that the commanded extrusion length is executed perfectly, in similar fashion to the other axes, since most FFF printers are base on an open-loop architecture (see Section 2.9). The work presented in Section 3.1 and Section 3.2 finds that the actual flow depends on liquefier temperature and feed speed. It is therefore of interest to find optimal parameters for the feed speed and temperature, for a desired track width. Additionally, an improved process model can improve the slicer model or the FFF design rules [185].

3.3.6 Stringing Length as a Response Measurand

Retraction is an important FFF technique if objects other than vases are desired. A vase print typically does not require a retraction, since it is one long continuous extrusion. Any other more intricate object requires that the extrusion is stopped and restarted after a nozzle repositioning, non-print move is completed.

The exit flow and retraction experiments in Section 3.2 found that both material oozing and stringing can occur during printing. FFF processing temperatures, apart from the retraction speed, will also play an important role. Both the nozzle temperature and object cooling fan speed will affect the extruder oozing and stringing. The temperature setting, however, needs to be optimised for the specific filament material to reduce thermal degradation, increase maximum possible feed speed as well as to optimise the post-extrusion solidification process (see Section 2.8).

3.3.7 Object Height as a Response Measurand

Object Z-errors and bed adhesion problems can arise if the print bed is not parallel with the horizontal plane generated by the nozzle movement relative to the build bed. Bed out-of-levelness can be caused by bending of the bed plate (see Section 2.3.2), incorrect mounting, thermal expansion

and dynamic effects. Here it is assumed that the bed is in a steady state, with small temperature fluctuations.

It is further assumed that the printer is already operational, in that an object can be printed without adhesion failure and that the basic millimetre to step conversion factors were set appropriately. The experiments performed with the object base height as a response measurand are aimed to test if the bed levelness can be deduced from the object height. The first three layers, which form the base height, were therefore printed with the same slicer settings.

The SPO method can, however, also be used to find an operational set of parameters. The optimal settings for bed adhesion were found by adjusting the *first layer height* and *first layer extrusion width* slicer parameters, using SPO and DOE.

A bed level test object was printed with different first layer settings and qualitatively compared to determine the influence of the two factors. The first SPO run and the verification run are shown in Fig. 3.31. Three SPO runs were required to find values for the two factors which resulted in significantly improved bed adhesion, independent of part position on the bed.

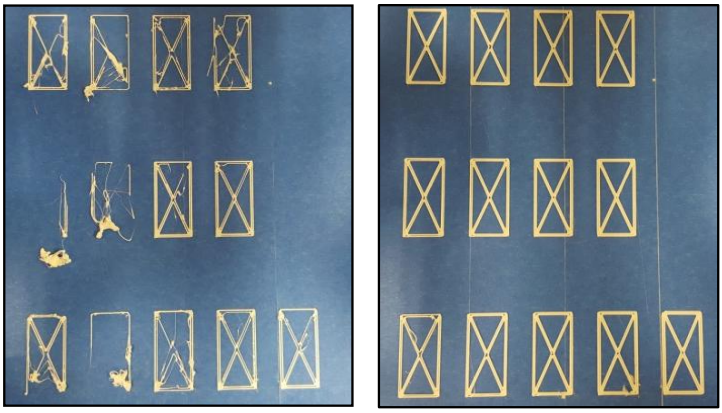


Figure 3.31: Bed adhesion optimisation, with the first Single Print Optimisation print run (left) and the verification print run (right).

A model was fitted and quickly showed that the initial assumptions for improving bed adhesion were wrong, by interpreting the signs of the fit coefficients in Eq. (3.11). It was first assumed that the *first layer height* (in millimetres) should be decreased from the previously used value. The idea was that this will force the track into the bed, therefore increasing the adhesion. Note that the *first layer extrusion width* is a percentage value, which modifies the first layer width (i.e. extrude more material).

Increasing the volumetric flow rate should increase track surface area, which will also increase the adherence force.

The tests found that the first layer height must be increased instead of decreased. A possible reason for this how the slicer uses this height parameter. Increasing the height, will also increase the amount of material, while decreasing it will decrease the volumetric flow rate. Increasing the height, however, also reduces the risk of printing in air instead of on the print bed.

The height was therefore increased from 0.2 mm to 0.35 mm, while the extrusion width was kept at 250 %. Note that it is therefore also important to consider how a certain parameter modifies the slicer behaviour.

3.3.8 Test Object Design

Increasing product complexity due to AM, necessitates *Design for Metrology*. The designer should not only satisfy functional requirements, but also consider how and why a feature must be measured to check if the functional requirements are met. This measurement requirement includes aspects such as traceability, cost-effectiveness, and uncertainty (see Section 1.5).

A standard test part for AM is proposed by Moylan *et al.* [186], see also [117, 165, 169] for more designs. These parts have a clear metrological design intent aimed at testing the ability and limitation of a machine as well as quantifying process errors. Included in the part design are many features, including holes, bosses and fine features in different orientations. Geometric dimensioning and tolerancing (GD&T) are also used to specify the measurement requirements and datums. This, however, requires multi-dimensional and expensive measurement instruments, for example a Coordinate Measuring Machine (CMM).

Furthermore, there are currently many test parts for FFF printers on the internet. They range from blocks, to more complex parts with for e.g. overhang features. Certain test parts are called *torture-tests*, intended to test the whole machine ability. Other test parts cannot be reproduced using the standard slicer algorithm. Examples of these are the acceleration fine-tuning test part presented in [72] or the ooze test part developed in Section 3.2.7.

Many of FFF test objects are, however, geared toward qualitative process inspection. A well-known torture-test object is the *3D Benchy* [187]. This part can test different process abilities and limits, but requires approximately 42 min to print with the centre-run settings used in this work. Quantitative 1D dimensional measurements are also possible, but the features were not clearly designed for such a purpose. This piece therefore gives a good indication of whether a printer is configured correctly or not, but it is not intended to be used for optimising a specific slicer parameter.

The test object designed for the SPO test is shown in Fig. 3.32 and can be measured with easy to use instruments. Furthermore, clear relationships between the CAD object, slicer parameter and the generated G-code can be deduced. The test part also only takes about 3 min and 24 s to print, with the centre-run settings. An input change of the selected slicer

parameter should lead to a measurable and predictable output change, if the slicer parameter is significant. This test object also does not require G-code modification and the slicer output is printable as is.

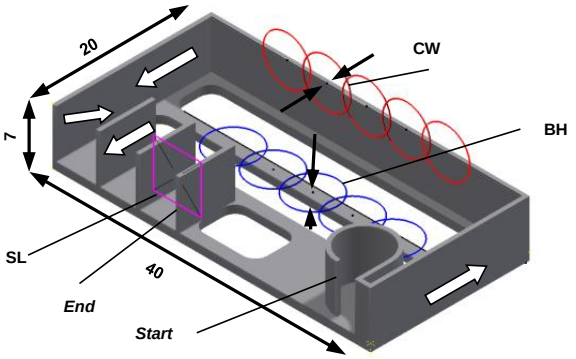


Figure 3.32: CAD of the design for metrology test object, with a height of 7 mm, length 40 mm and a width 20 mm. The main features are the three short walls, one long wall, a semi-circle and a base layer, which has a thin road between the two shorter lengths of the long wall. The red circles indicate the measurement positions for the centre width (CW), blue circles for base height (BH) and the magenta square area for the stringing length (SL) photos. Large arrows indicate the print direction from start to end.

The part design ensures that the slicer generates G-code, which extrudes one long road without stopping, i.e. a long, thin, external perimeter wall. The wall has a CAD defined width of 0.5 mm, which is smaller than the slicer single perimeter width (0.72 mm). This ensures that the wall is sliced as a single track. The width of this track is measured along the 40 mm long wall to determine the *centre width* response measurand. Note that the length of the wall is critical to ensure that the printer is not accelerating or decelerating over the measurement section.

The part base is not a solid rectangle, but several slots are cut into it. This significantly reduces the print time, since the infill volume for a 0.3 mm layer height is 240 mm^3, whilst that of the perimeter wall is 17 mm^3. The base track that is parallel to the long wall is used to measure the base height. The wall also has a smaller inner flange to support the measurement of the total height as well as reduce the risk of bed adhesion failure.

The designated measurement instruments used are shown in Fig. 3.33. The digital micrometer and the USB microscope are examples of commonly available and easy to use dimensional measurement instruments. The micrometer has a range of 25 mm and a manufacturer specified uncertainty of ±2 µm.

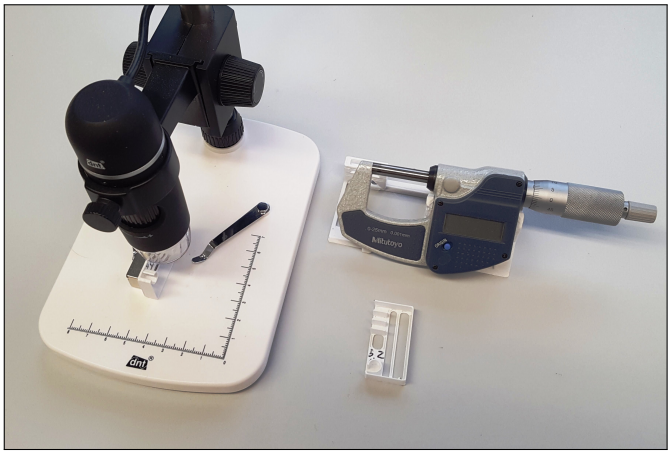


Figure 3.33: The entry level digital microscope, test piece and the micrometer, mounted on a 3D-printed stand, which were used for the measurements.

Three short walls are used to test the retraction effectiveness and the measurable feature is the stringing length, which is photographed with the digital microscope. The stringing length is qualitatively categorised into groups, ranging from no stringing visible to long strings. Note that the digital microscope can also be used to measure the track width, but will only measure the top, or last layer.

The micrometer measures a maximum width over the face of the 6 mm diameter anvil. The compressive force exerted by the micrometer is limited by a ratchet, further increasing the repeatability of the measurement.

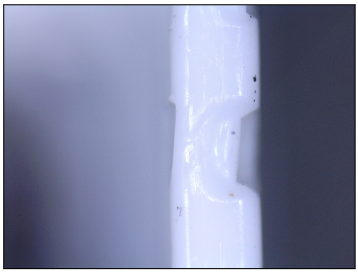
Another option to measure the track width is to use a vernier calliper, which is also cheaper, but less accurate. One risk of a using vernier calliper is that the structure can be cut by the knife edges, as shown in Fig. 3.34.

This was achieved only when a single layer was pinched by the vernier. Measurement over many layers, however, did not produce any visible damage nor did micrometer measurements visibly damage the tracks.

Four G-code lines are taken from the file for a centre parameter object at DOE run position 0 and shown in Listing 3.5.

The first line is part of the file preamble and comments that the slicer expected extrusion width is 720 µm. The next line is the start of the perimeter wall extrusion. It defines a feed speed for the subsequent extrusions at 2700 mm/min. The second last line defines the start position of the last line, which is the long side of the wall, which will be measured to determine the centre width.

The last two lines in Listing 3.5 and the expected width equation for an oblong track shape (Eq. (2.63)) are used to calculate the expected track width as 719 µm, with the layer height h, track length L_{xy}, fila-



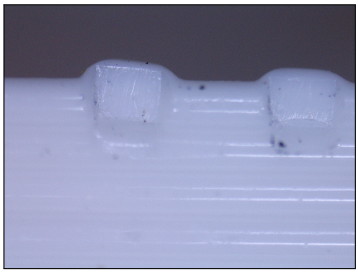


Figure 3.34: Photos of the damage done to a single track by using a vernier calliper. Top view (left) and side view (right). Note that no apparent damage was found when measuring over many layers and that this occurred when measuring only the top most layer (single track).

Listing 3.5: G-code command file excerpt.

```
; perimeters extrusion width = 0.72mm
...
G1 X21.022 Y30.335 E2.01032 F2700.000
...
G1 X39.750 Y25.500 E3.24077
G1 X39.750 Y64.500 E4.44114
```

ment diameter D_f and the extrusion length E_f equal to 0.3 mm, 39 mm, 2.85 mm and 1.200 37 mm.

The effect of filament diameter and track height variation on the track width can be estimated. A diameter variation of ±100 µm results in a range of 674 – 766 µm or ±49 µm for the track width. The diameter change therefore represents a 4 % variation.

Interestingly, the same percentage of variation of the layer height (289 – 311 µm) results in ±43 µm range for this specific extrusion.

The diameter of the filament is expected to vary in this order of magnitude according to manufacture supplied information, but it is not expected that the layer height variation is this large, since the micro-stepping resolution of the Z-axis is 0.4 µm (see Section 2.4.3).

3.3.9 SPO Test Object Experimental Setup

The results of the centre width, stringing length and the base height measurements are presented next. These results are used to define a basic cost model, from which an operation point is selected and verified. Four print runs were executed, namely *Test A*, *Test B*, *Test V* and *Test R*.

Test A is a central run where all the objects were printed with the centre values of all the parameters. Test R is a reverse run of Test A where only the positions are reversed. Test B is the DOE run table experiments and a five object verification run, Test V, was used to test the predicted values of the fitted models, at a selected operating point.

The object base height and the centre width were measured with the micrometer. Five measurements were taken over the length of each feature and the mean was calculated. The average standard deviation for the base height was 7 µm and for the centre width was 6 µm, indicating that the instrument is fit for the task and that there are no significant deformations along the tracks.

3.3.10 SPO Centre Width Results

The mean centre width for Test A, Test B as well as the centre width prediction by the model (Fit) are shown for each experiment in Fig. 3.35. The average of the centre width for Test A is 632 µm, with a standard deviation of 14 µm. The centre width standard deviation is reduced to 7 µm, by removing the first two runs, which appear to be outliers. This standard deviation is comparable to the average standard deviation of the feature measurements.

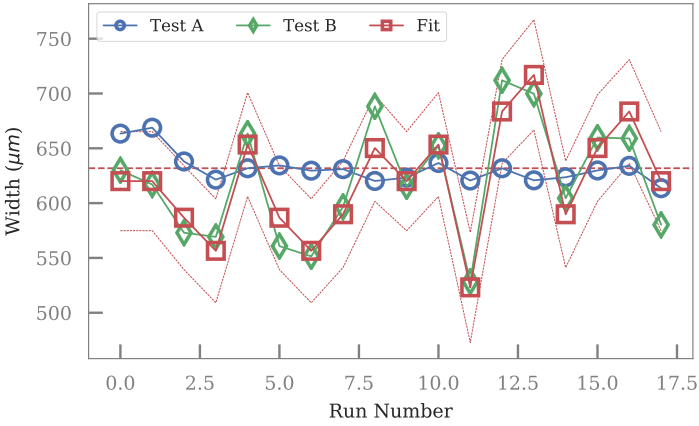


Figure 3.35: Measured centre width (CW) for each run, for Test A and B as well as the predicted width values from the fit. Unfilled circles are Test A, diamonds Test B and squares for the fit. The horizontal dashed lined indicates the mean CW of Test A. Note the first two runs of Test A seem to be outliers.

The average of Test A is similar to that of Test B at 620 µm, as expected with a balanced DOE design. The standard deviation is larger and equal

to 54 µm. Both Test A and Test B do not show a significant linear trend, which could have indicated drift. This provides evidence that the process is reproducible and stable.

The expected centre width is 719 µm and the mean value for Test A is 87 µm less than this value. The prediction is that the slippage of the filament feed drive increases, with increasing speed or decreasing temperature, since the force required to generate enough pressure to extrude the molten material increases with higher viscosity or faster speed (see Sections 2.8.9 and 3.1).

This prediction is supported by looking at the influence of each experimental factor on the centre width, as shown in Fig. 3.36 (A to E). The perimeter-speed factor shows a linear trend, where increasing speed decreases the track width and the temperature plot only indicates a possible increase of centre width at a higher temperature.

Retract speed and object position (cx and cy) do not indicate a clear relationship, as the centre width seems to be spread equally around the mean value.

This is confirmed by the OLS regression with iterative factor removal, which finds a linear model (see Table 3.9). The factor coefficients are of the same order magnitude, with perimeter-speed ($x0$) having a larger negative coefficient and temperature a smaller positive coefficient ($x1$). The other factors, quadratic terms and interaction effects, are found not to be significant.

Table 3.8: OLS fitting statistics for the centre width (CW) model.

Statistic	Test B: CW
Residual mean	1.9×10^{-13}
Residual std	19.386
R-squared	0.871
Adj. R-squared	0.853
F-statistic	50.52
Prob (F-statistic)	2.2×10^{-5}
Cond. No	1.50

The test statistics for the centre width OLS regression are shown in Table 3.8. The R-squared and Adjusted R-Squared values are sufficiently close to one and the F-test indicates that the fit is significant. The filament diameter is assumed to be constant here, but this can also be a significant factor. This effect is not considered in the model, but can influence the model accuracy.

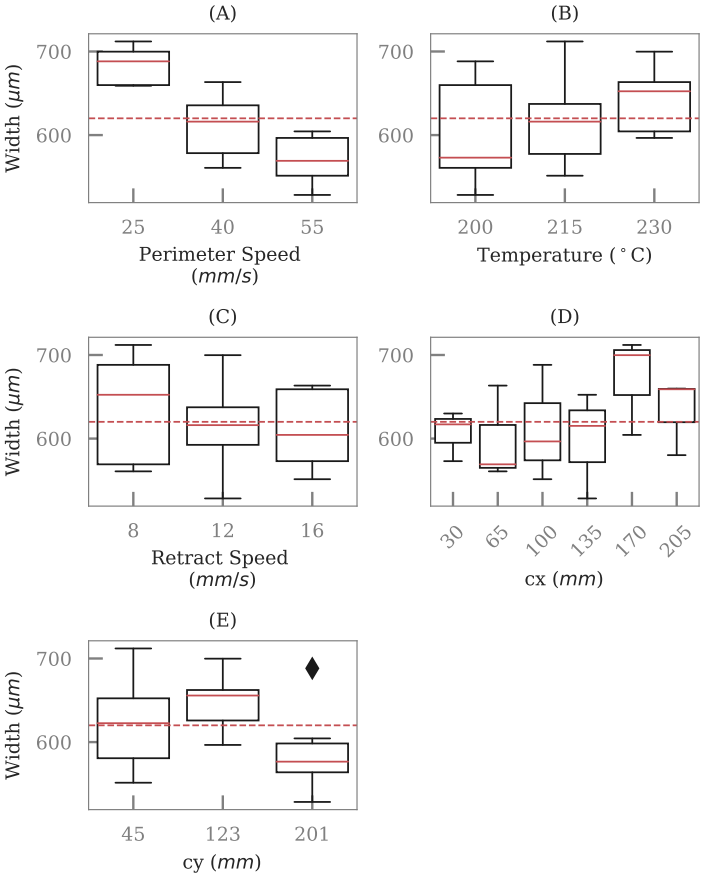


Figure 3.36: Centre width (CW) grouped by each factor. Boxplots show the median value and the spread of CW for the factors, diamonds indicate outliers and the dashed horizontal line the mean CW.

Table 3.9: Response variable factor coefficients for the centre width (CW) model (x0: Perimeter-Speed; x1: Temperature).

Y	Factor	coef	std err	t	P >|t|
CW	Intercept	620.0444	4.864	127.467	0.000
	x0	-63.5500	6.661	-9.541	0.000
	x1	33.3500	6.661	5.007	0.000

3.3.11 Stringing Length

Photos of the end of the last short wall were taken to evaluate the stringing length (see Fig. 3.37). The slicer started each layer at the semi-circle feature, then completed the long wall, after which it produced the three short walls. Each short wall was started on the inside of the part and ended at the object edge. The last short wall is closest to the semi-circle feature and not only requires a retraction after extrusion completion, but also a change in the Z-axis position upwards to start a new layer or object.

The last short wall of the test object was chosen for the stringing length quantification, since the other two short walls did not show significant stringing effects. The additional vertical move is assumed to be the cause for this difference between the three walls.

A small amount of ooze material can be pulled out of the nozzle, if it is first moved upward, by the adhesion force to the already deposited track. In contrast, the horizontal move reduces the chance of material stringing, since the fine string will be touching the hot nozzle, which can sever the string (see inset in Fig. 3.37).

It is expected that stringing length should increase with decreased cooling time or increased temperature, since the track will be less solidified. A faster retract speed will decrease the delay time, before the nozzle moves to a new location, increasing the stringing of material. On the other hand a faster retract speed may also reduce the liquefier pressure quicker, reducing the amount of material oozed (see Section 3.2).

The stringing length for each run was grouped into one of four levels (0 to 3), with zero indicating no stringing and 4 the worst case. Vertical lines are used to differentiate between Level 2 and Level 3 in figure Fig. 3.37.

The same regression method was used for the stringing length results for finding the significant factors. The test statistics for the stringing length OLS regression are given in Table 3.10 and are similar to that of the centre width statistics. The response variable is, however, only integer levels, which can reduce the R-squared value.

The temperature ($x1$) and retract-speed ($x2$) are found as significant factors in Table 3.11. The p-value of the retract-speed is slightly outside the 95 % confidence level at 0.057.

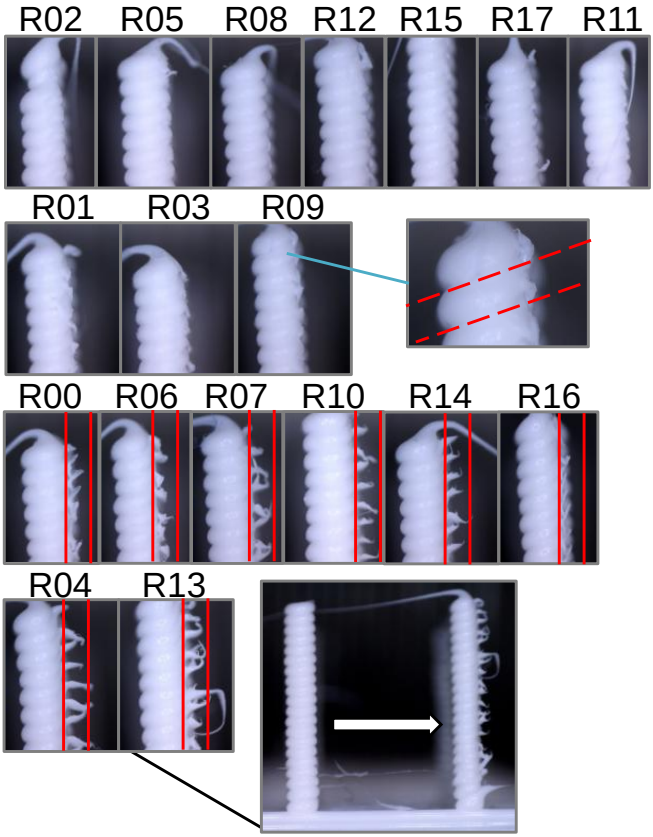


Figure 3.37: The stringing length (SL) for each run in Test B with the run number indicated on top. Photos with approximately the same SL are grouped in rows, sorted with increasing SL. The mean track width, approximately 620 µm, indicates scale. The first insert (second row) shows the track vertical deformation and the second inset (last row) compares the SL between the last two short walls of Run 13.

Table 3.10: Stringing length (SL) OLS fitting statistics.

Statistic	Test B: SL
Residual mean	1.5×10^{-16}
Residual std	0.529
R-squared	0.768
Adj. R-squared	0.737
F-statistic	24.87
Prob (F-statistic)	1.7×10^{-5}
Cond. No	1.50

Table 3.11: Response variable factor coefficients for the stringing length (SL) model (x1: Temperature; x2: Retract-Speed).

Y	**Factor**	**coef**	**std err**	**t**	**P >|t|**
SL	Intercept	1.1667	0.133	8.796	0.000
	x1	1.1250	0.182	6.194	0.000
	x2	0.3750	0.182	2.065	0.057

3.3.12 Base Height

The base height measurements are presented in Fig. 3.38 for each run. The G-code commanded Z-position is 950 µm and the run plot indicates a significant deviation from this value. The mean base height for Test B is 886 µm, which is 64 µm less than the set point. The maximum deviation is however −155 µm. This is more than half the standard layer height of 300 µm.

The base heights for Test A and Test B agree well. The average difference between the tests is 1 µm and the range, maximum minus minimum, is 28 µm, with a standard deviation of 8 µm. This standard deviation is comparable to the average standard deviation of the feature measurements, which supports the assumption that the process is reproducible and under control.

The two additional plots in Fig. 3.39 group base height versus the X and Y object positions (cx and cy). The dashed horizontal line indicates the mean base height.

A clear trend is visible and is most likely due to either a time based factor or the position of the objects. The base heights grouped by position provide further insight into this trend. The X-positions are unfortunately not randomised and increase with run number. Additional data is there-

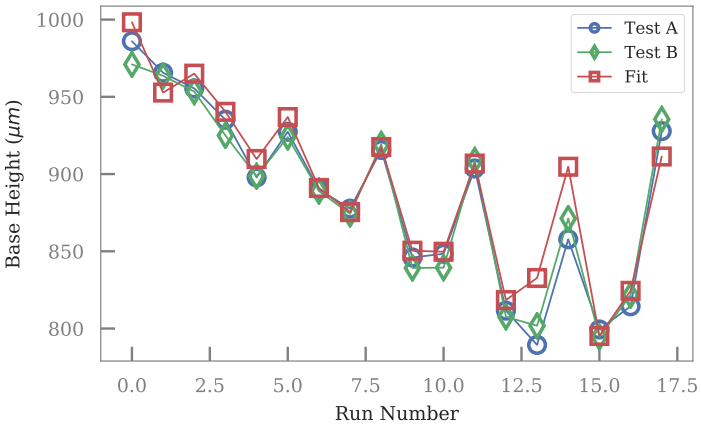


Figure 3.38: Base height (BH) measurements for Test A and Test B, as well as the fitted values, plotted against the run number. The reproducibility between Test A and Test B is high and the fit seems to predict the BH values.

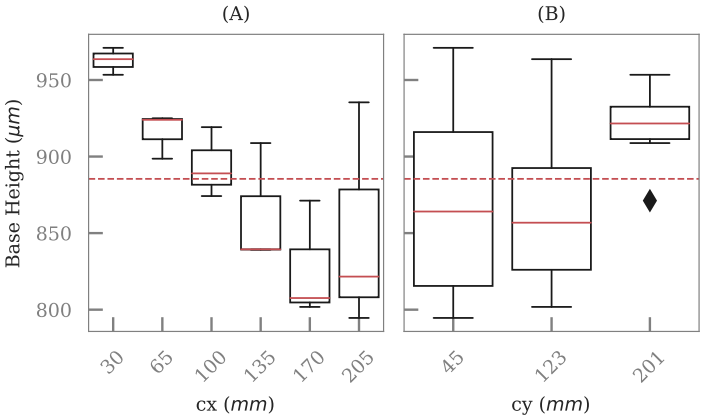


Figure 3.39: Boxplots of base height grouped by object X and Y position (cx and cy) in sub-plot A and B respectively. The mean base height is indicated with the horizontal dashed line. Factor cx can be a significant factor. Diamonds indicate possible outliers.

fore required and a reverse centre run, Test R, was performed. The object positions were reversed, so that the last object position of Test A is the first printed object position of Test R.

The difference between Test A and Test R (plotted with a reversed run number) is shown in Fig. 3.40. The difference between Test A and B is shown for reference, as well the difference between the predicted values for Test R, using the model fitted to Test B. This shows a linearly decreasing base height difference, as time increases.

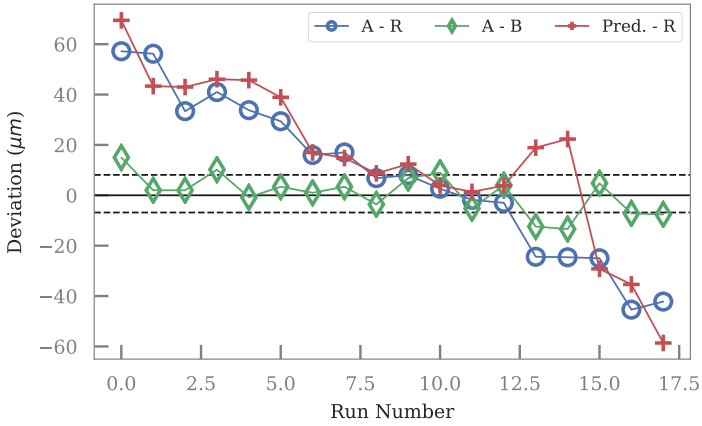


Figure 3.40: Unfilled blue circles show the BH (Base Height) difference between Test A and Test R (run order reversed) and indicate a linear trend. Green unfilled diamonds compare Test A and Test B, while the dashed horizontal lines is the standard deviation of $A - B$. Red pluses are the difference between Test R and the predicted Test R values with the Test B model.

It is assumed that the object base height deviation is caused by the parallelism error between the nozzle plane and the build bed. A possible reason for the apparent trend in the difference between the forward and reverse measurement is the mounting of the build bed as discussed in Section 2.3. Any pitch or roll around the Y-axis will influence the bed level. Other factors include thermal expansion effects.

An OLS regression as well as a robust linear models (RLM) fit were applied. The RLM method was used to reduce the possible influence of outliers, since OLS is sensitive to outliers and the base height grouped by *cy* indicated a possible outlier.

Test statistics provide further insight, as reported in Table 3.12, since the two main input factors seem to be highly correlated, as indicated by the high condition number. This is called multicollinearity and suggest that the predictors do not provide unique or independent information

for the regression analysis. The effect of this on value prediction can be reduced by performing the experiment at the same location.

Table 3.12: OLS and RLM fitting statistics for the base height (BH).

Test B: BH		
Method	OLS	RLM
Residual mean	-1.1×10^{-11}	−7.944
Residual std	13.558	14.167
R-squared	0.944	
Adj. R-squared	0.920	
F-statistic	40.17	
Prob (F-statistic)	4.4×10^{-7}	
Condition Number	2.5×10^{5}	

Further research into this effect is, however, currently limited by the focus of this work, which only requires a good enough fit to evaluate the concept of bed level prediction and correction with test object height measurement.

A quadratic fit of the base height with RLM, using *cx* and *cy* as input factors, is used here to statically correct for the bed level deviation. As discussed, the model does not capture all the characteristics of the data, but was nevertheless used to see if it can optimise the FFF process.

The model factors and coefficients are given in Table 3.13. The p-values for all the factors are less than 0.001, indicating significance and coefficients for *cx* and *cy* are the largest and negative. The coefficients of the interaction factor *cx:cy* and the quadratic terms are smaller, but of equal order of magnitude relative to each other, and positive.

Table 3.13: Response variable factor coefficients for the base height (BH).

Y	**Factor**	**coef**	**std err**	**z**	**P >|z|**
BH	Intercept	1114.782	23.206	48.04	0.000
	cx	-2.2393	0.276	-8.099	0.000
	cy	-1.5449	0.284	-5.446	0.000
	cx^2	0.0035	0.001	3.565	0.000
	cy^2	0.0047	0.001	4.481	0.000
	cx:cy	0.0055	0.001	6.862	0.000

The base height model is tested here to see if it can be used to correct for the bed level error by modifying the Z-position of each object with Eq. (3.12). This is done by supplying a Z-offset, dZ, for each object when it is sliced. The expected G-code base height, H_{gcode}, is deducted from the predicted base height, $\hat{H}_{pred}$, and added as a Z-offset to each object. This changes the Z position of each part. Note that this is not the same as supplying the firmware with a Z-offset lookup table.

$$\begin{aligned} \hat{H}_{pred} &= \beta_0 + \beta_1 cx + \beta_2 cy + \beta_3 cx^2 + \beta_5 cy^2 + \beta_5 cycx \\ dZ &= H_{gcode} - \hat{H}_{pred} \end{aligned} \tag{3.12}$$

3.3.13 Cost Model

The last step in this work is to validate the experimental models with a verification print. This is done by selecting an operation point and comparing the predicted results with the measured.

The cost for the centre width (C_{CW}) response factor is defined here as the absolute deviation from the expected slicer model track width $W_{expected}$, as calculated from the G-code command, from the predicted track width, W_{model}. The cost of the stringing length (C_{SL}) response factor is defined as increasing with increasing stringing length level. Both costs are normalised to form $C_{norm:Y}$ and added together, each with a weight of 0.5, to calculate the total cost, C_{total} with Eq. (3.13).

$$\begin{aligned} \Delta Y_{CW} &= \left| W_{expected} - W_{model} \right| \\ C_{norm:Y} &= \frac{\Delta Y - min\left(\Delta Y\right)}{range\left(\Delta Y\right)} \\ C_{total} &= 0.5 C_{CW} + 0.5 C_{SL} \end{aligned} \tag{3.13}$$

The contour plots in Fig. 3.41 show the individual costs for the both the centre width and the stringing length over the DOE parameter range.

The optimal point within the parameter range is at either the respective minimum or maximum, since both relationships are predicted to be linear. The centre width model requires the maximum temperature, while the stringing length optimum point is at the minimum temperature. A compromise between the two must therefore be reached or an additional cost factor should be considered.

Alternative ooze reductions strategies, which are not considered here, are the optimisation of the nozzle tip (see Section 2.8.6), improvement of the cooling fan, the use of an external wall around the print object or special retract moves (e.g. first moving horizontally, then vertically).

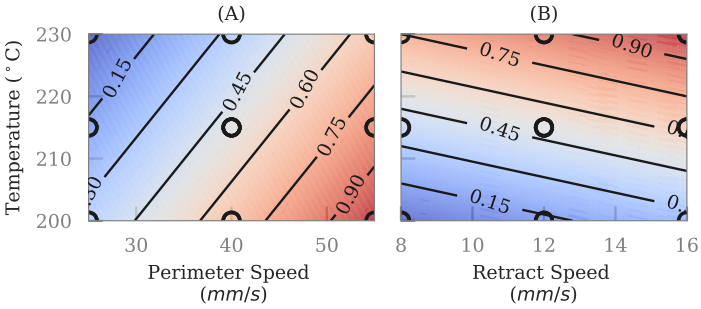


Figure 3.41: Normalised cost of (A) the centre width (CW) and (B) the stringing length (SL), against the model factors for each response factor. The unfilled circles indicate DOE run locations.

Printing speed is introduced as a third cost factor here and the fastest perimeter-speed of 55 mm/s is selected. The retraction-speed will have relatively small influence on the total printing speed and the slowest speed of 8 mm/s is chosen for this. The minimum cost point is then found and used as the operating point. The resulting costs for centre width, stringing length and the combined total cost are shown in Fig. 3.42.

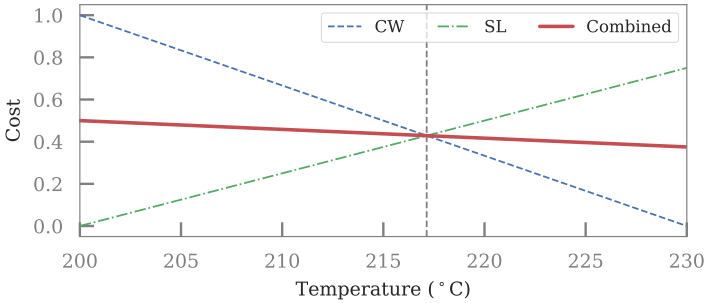


Figure 3.42: Optimal point selection by comparing the different costs.

Two optimal points can be seen in Fig. 3.42: one where the cost of the centre width and the stringing length crosses and the other where the combined cost is minimised. The crossing point is interesting, since it is close to the current supplier recommended values. The total cost is, however, a minimum at 230 °C and this point is therefore selected.

3.3.14 Verification Prints

After the selection of the operation point a five object verification run was performed (Test V). Four objects were printed in the corners of the print bed and the fifth in the bed centre. The order of printing was the same as with Test A and Test B: starting at the XY-origin, first increasing in the Y-axis and then in the X-axis. A Z-offset was added to each object to adjust for bed level misalignment and all the objects were sliced with the selected operating point parameters.

Photos of the stringing length of the verification run are shown in Fig. 3.43 and are similar to Level 2 of Test B (see Fig. 3.37). The stringing length model also predicts a stringing length level of two. The verification run therefore confirms this prediction.

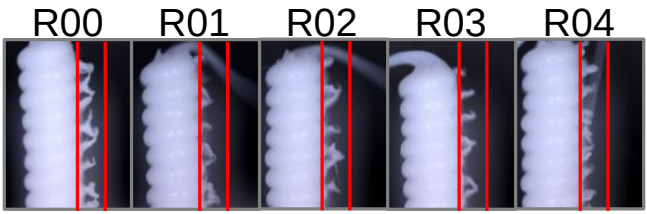


Figure 3.43: The stringing length (SL) of Test V, photos are taken from the same position as with Test B. All SL are qualitatively similar to Level 2, as defined in the Test B SL, results

The predicted centre width at the operating point is 590 µm, while the predicted centre width, before modelling for all slicer parameters, is 719 µm. The measured centre width mean is 630 µm, which is 40 µm more than the predicted centre width. This is, however, an improvement from the pre-modelling deviation of 89 µm, improving the prediction with 55 %.

The centre width deviation is shown in Fig. 3.44. This shows the deviation of Test A and Test B from the 719 µm centre width as well as the verification Test V difference from the predicted 590 µm.

Even though it was found that the base height model does not sufficiently describe the measured data, the verification plot in Fig. 3.45 finds that the individual bed Z adjustments do improve the dimensional accuracy of the final product. Note that the centre object in Test V was also printed at an unique XY position.

The expected base height is 950 µm. The median, as shown by the horizontal line inside the boxes for the verification print in Fig. 3.45, is only marginally closer to the ideal value, while the mean base height deviation decreases from −65 µm to 25 µm. The experimental results indicated far better reproducibility. It is therefore predicted that an improved model and possibly an extension of the DOE plan should result in a more accurate prediction and bed level correction.

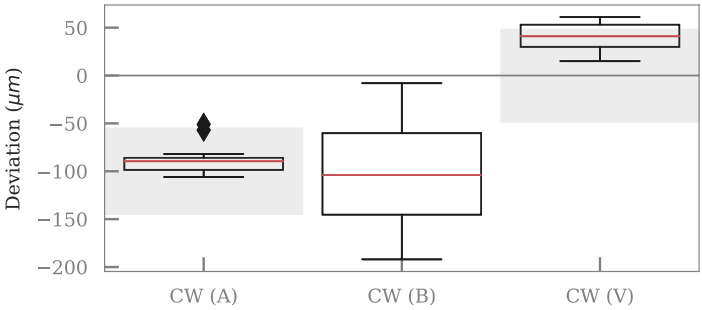


Figure 3.44: The deviation from expected centre width (CW) for Test A, B and V with boxplots CW (A), CW (B) and CW (V) respectfully. CW (A) and CW (B) are the deviation from the G-code derived width and CW (V) from the OLS model predicted width. Shaded areas show predicted confidence intervals for point observations.

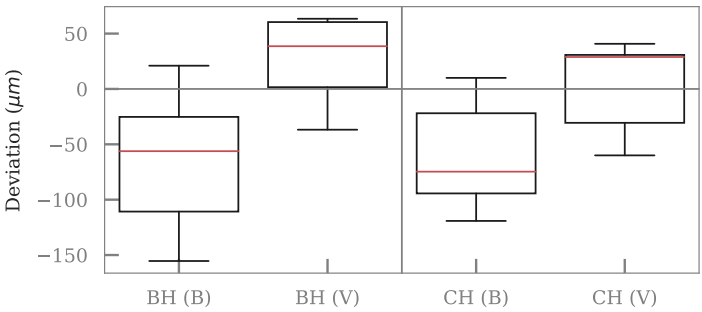


Figure 3.45: The deviation from the G-code defined part height for the base height (BH) and centre height (CH) for Test B and Test V.

The total height at the centre of the long wall (CH) of the objects were also measured, with five points along the feature. The CAD of the object specifies a height of 7 mm, whilst the G-code set the height at 6.950 mm, with a constant layer height of 0.3 mm and a first layer of 0.35 mm. Interestingly the mean of the measured CH deviation reduces from −61 µm, for Test B, to 2 µm, for Test V.

3.3.15 Discussion

The application Single Print Optimisation (SPO) for the optimising of slicer parameters to improve dimensional printed part accuracy and reduce the expected result uncertainty, by Ordinary Least Squares (OLS) regression and Design of Experiments (DOE) was shown. Furthermore, a test part, which is fast to print, requiring approximately only three and a half minutes, is presented. This test part is designed to be measured with accessible and easy-to-use, yet sufficiently accurate, measurement instruments.

The key process in FFF is the deposition of a single track, this includes the start, middle section and the end of the track. This should be done as fast, but also as reliable and reproducible, as possible.

The track width and height determine the produced part geometric accuracy, where the track width is varied by slicer, to achieve certain functions, such as bonding between adjacent tracks, fast infill, bridges or fine features, while the height is normally kept constant.

The track width is a function of the amount of material fed into the liquefier during the print move and this amount can vary due to the filament feed mechanism slippage. Note that the relative print speed and track height can also influence the track shape, and therefore the track width, as discussed in Section 2.8.14. This is, however, not considered here.

The centre width measurement results find that the slippage depends on the slicer settings for temperature and perimeter-speed. A DOE strategy was used to find an empirical model for this relationship, using OLS regression.

A verification print was then performed and found to improve the print result, by reducing the error in expected track width with 55 %. The verification print was also tested against the prediction of the model and found to be sufficient, but with room for further improvement. This is considered in more detail in Section 3.4, where the in-process measurements are compared to the post-process width measurements.

Not only was the slippage effect modelled, but the bed level was also estimated with object height measurements. Another model was fitted and the Z starting positions of the individual verification print objects were adjusted to compensate for the out of levelness of the printing bed. It was found to reduce the mean object height error from −61 µm to 2 µm.

Stringing, as result of non-print moves, was also modelled and found to depend on the temperature and the retraction speed, where a longer solidification time or a lower temperature, reduces the stringing effect. The factor reduction method finds that temperature is the most significant

input factor, since it has the largest coefficient. This is also contrary to the expected result. The verification print was performed at the selected optimal point and agreed well with the predicted values.

The retraction tests performed in Section 3.2.6 only varied the retraction speed and length. These tests also found that the speed is not as significant, which agrees with the SPO DOE results found here.

This work shows that it is therefore possible to optimise FFF slicer settings, with a single print run, by performing multiple experiments sequentially, where each experiment is a test object sliced with different parameters.

Additionally, DOE run tables can also be merged, if the response variables have sufficiently independent input factors. The principle can be expanded with more control factors, response variables, as well as more complex optimisation methods and models. An example is to include the filament width as an additional factor. Dynamic effects can also be modelled, in order to find optimal acceleration parameters.

3.4 In-Process Monitoring and Modelling Results

This section combines the in-process and the post-process measurements, obtained during the Single Print Optimisation test run in Section 3.3. The in-process data were collected with the speed, filament width and pressure measurement methods presented in Sections 3.1 and 3.2.

The processing of the raw measured data is presented first, followed by the results for each experiment, where an experiment is a printed object in the DOE run table. The results are fitted to the Bellini pressure drop model, as well as an empirical model, after which different implementation strategies are discussed.

3.4.1 Data Processing

Real time capturing of the feed speeds (gear and filament) and the filament diameter were presented in Section 3.1. The liquefier pressure measurement is explained in Section 3.2, while the printer also reports the extruder temperatures. This means that there are at least five variables which are monitored during printing. The vision and printer data streams are, however, sampled at different frequencies and reported asynchronously to the main application.

The main application coordinates the whole process and gathers the data during the print run. It stores the data in a tabular file format, after the print run is completed, where each row in the file represents a time point. Values that are not available at a specific time point are marked as empty.

Off-line processing is then used to shift and interpolate the data file. This creates a single dataset which describes the entire print run, in terms of time and the five variables. Other variables are also stored, such as the frame number, the temperature PWM control level and the bed temperature, but are not discussed here.

Another data stream is introduced next, which is the G-code command file. The original print run commands are parsed and an extrusion (feed) speed profile is generated. Note that speed ramping was not considered in the generation of the commanded feed speed profile. The command line number and text is linked to each linear section of the feed speed profile. This is very useful, since the G-code command can be now be linked to the measured data. That is, if the commanded speed is aligned with the measure data.

The commanded profile must be aligned to the measured speed data, since the G-code is unaware of the heating and cooling delay times, as well as the acceleration or deceleration time implemented by the firmware (see Section 2.6). Note that this will only shift and stretch the start and end time points of each linear speed command. The alignment becomes critical for long prints, since the aliment errors will, in general, accumulate, which shifts the measured and commanded speeds away from each other.

The total print time for the SPO *Test B*, for example, is approximately 42 min, which result in 75 600 sample points (at 30 FPS), where each sample contains the interpolated values for all the measurements.

The details of the G-code command alignment to the measured gear speed is not presented here. It does, however, involves several non-trivial steps, including parsing the G-code file, calculating the commanded feed speed, filtering the measured data to represent a step function and aligning each G-code command position to the measured time frame. The final result is a third data stream containing information of the object number, line number, commanded speed and a flag indicating if a certain time point is part of an *interest line* group.

The interest lines, in this case, is the long wall section which was used to measure the track width of the test object (see Section 3.3.8). This allows for the in-process measured data to be compared to the post-process measured data.

An example of the processed data for the SPO *Test B* is shown in Fig. 3.46. The commanded, gear and filament speeds, as well as the measured pressure, are shown. Note that the Y-axis is clipped to increase visibility. Sharp peaks indicate retractions or refill extrusions and flat segments print extrusions, while gaps indicate non-print moves or heating delay times.

Unfilled circles show the points selected at the predefined interest command lines. It can be seen that these points are approximately on the long flat segments. These sections included the wall sections before and after the section used for the centre width measurement. The time location of three short walls are indicated by the three peaks. Lastly, the short speed dip, before the flat section, represents the semi-circular feature, where the firmware presumably reduces the speed to enable cornering (the arc is made of several short linear extrusions, see also Section 2.6.1).

The detail of a single interest line is shown in Fig. 3.47, and includes the temperature data along with the filament width, gear speed and the pressure.

The filament speed is not shown here for clarity, while the solid horizontal line indicates the mean value for all the interest lines of a specific

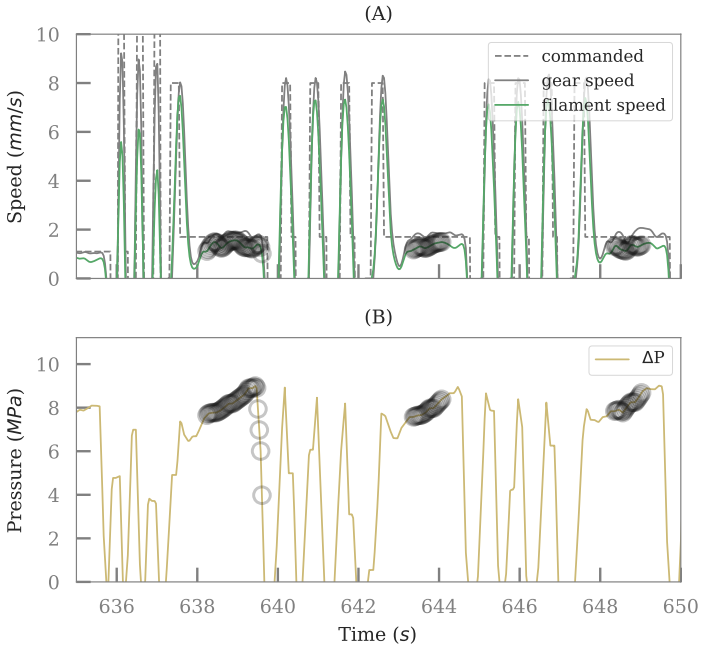


Figure 3.46: In-process data measured for (A) gear and filament speed, compared with the commanded speed (dashed line), and (B) the measured pressure. Unfilled circles indicate the selected points for the *interest line*.

object. The measured temperature was found to be approximately equal to the G-code commanded level and the commanded temperature level is therefore used throughout the rest of this section.

Some interest line points do not fall on the flat section (see for example the first pressure interest line segment in Fig. 3.46). A mean for each measured variable over all the interest lines for each object was therefore calculated. This mean, along with all the selected interest line points for a single object is shown in Fig. 3.48, for the speed, filament diameter and the pressure variables. The standard deviation is also shown with dashed lines.

The mean gear speed for this object is (1.81 ± 0.11) mm/s and the mean filament speed is (1.33 ± 0.10) mm/s, while the commanded speed is 1.7 mm/s.

The uncertainty for a 2 mm/s speed was estimated as ± 0.06 mm/s (see Section 3.1.8). Significantly, the standard deviations for this object are of the same order of magnitude as the estimated uncertainty.

The mean gear speed is slightly faster than the commanded interest line

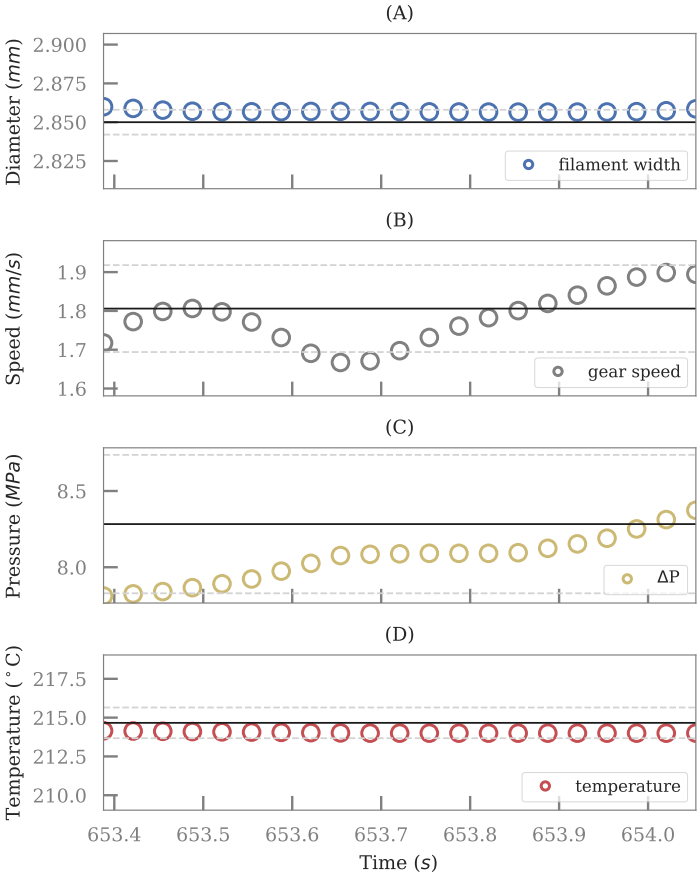


Figure 3.47: Data points for a single interest line of an object, with (A) filament diameter, (B) gear speed, (C) pressure and (D) measured liquefier temperature. Solid and dashed horizontal lines indicate the mean and standard deviation of all the interest lines of the specific object.

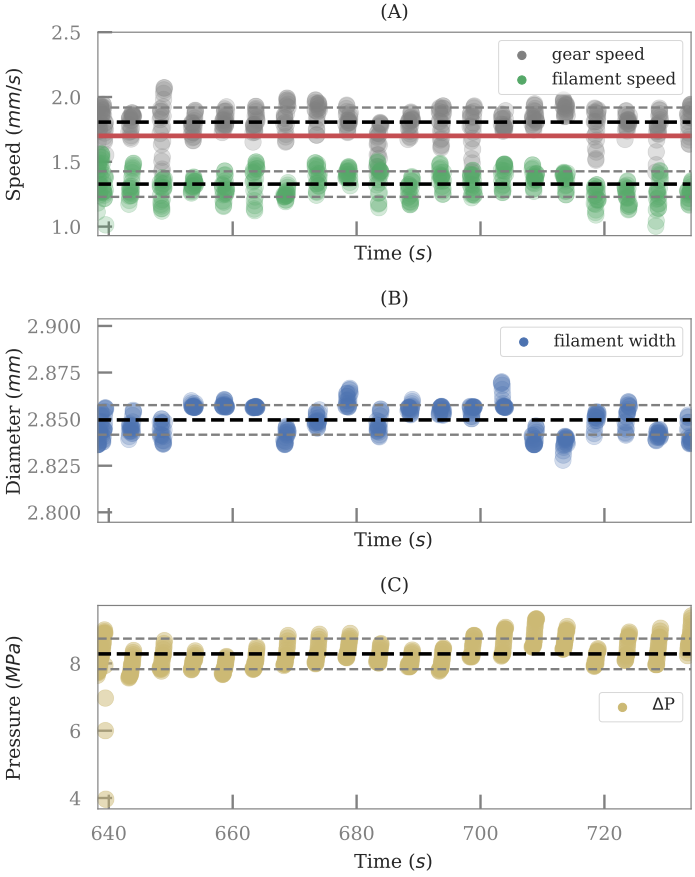


Figure 3.48: Data points for all the interest lines of a specific object in the print run, with (A) the measured speeds, (B) the filament width and (C) the pressure. Solid and dashed horizontal lines indicate the mean and standard deviation. Thicker dashed lines in (A) is the means of the gear and filament speeds.

feed gear speed, while the filament speed is slower. Note that this data is for a single object in the print run. The standard deviation represents the measurement and process variability combined. In this work it is assumed that both variabilities are independent and constant. Local variation of the filament diameter, however, will affect the pressure and therefore the filament speed, but this effect is neglected and only the mean value of each variable per object is used.

The mean filament width for this object is (2.85 ± 0.01) mm, which agrees well with the manufactured specification of (2.85 ± 0.10) mm. The measured standard deviation is equal to the width resolution limit ± 10 µm (see Section 3.1.7).

A mean liquefier pressure of (8.28 ± 0.45) MPa is calculated for this object and the standard deviation is larger than the estimated uncertainty of ± 0.26 MPa (see Section 3.2.4).

The measured values are further used to determine derived quantities, such as the feed slippage (Eq. (3.7)), predicted track width (Eq. (2.63)) and the input flow rate ($Q_{in} = A_f v_f$), where A_f is filament cross section area and v_f is the filament speed into the liquefier. The derived values for the example object are shown in Fig. 3.49.

The mean feed slippage percentage for this experiment is (-26 ± 5) %, the mean predicted track width is (0.58 ± 0.04) mm and the mean input volumetric flow rate is (8.47 ± 0.64) mm^3/s. The feed slippage standard deviation of ± 5 % agrees well with the ± 5 % half width limits of the speed range found in the free air feed slippage extrusion experiments (see Section 3.1.9)

3.4.2 Combined Test Object Results

The measured and derived data were collected for the interest lines of all the test objects printed in *Test B*, where each object was printed with slicer parameters according to the SPO DOE plan (see Table 3.6). The results are shown in Fig. 3.50 for the filament diameter, feed speeds and the pressure.

A prediction of the expected track width can be made if both the track shape and the steady state input flow are known. The oblong shape is used here to predict the average track width of the interest lines for each object (see Section 2.8.14 and Eq. (2.63)).

The measured centre width is compared to the estimated track width in Fig. 3.51, with error bars for the predicted width equal to the measured standard deviation for each object and error bars for the measured track width represents ± 10 µm.

The predicted mean centre track width is (541 ± 21) µm, while the measured mean centre width is (620 ± 6) µm (see Section 3.3.10). This difference can be explained by the simulation work by Comminal *et al.* [146], who predicts that the track shape will deviate from the oblong form, depending on the print speed and layer height.

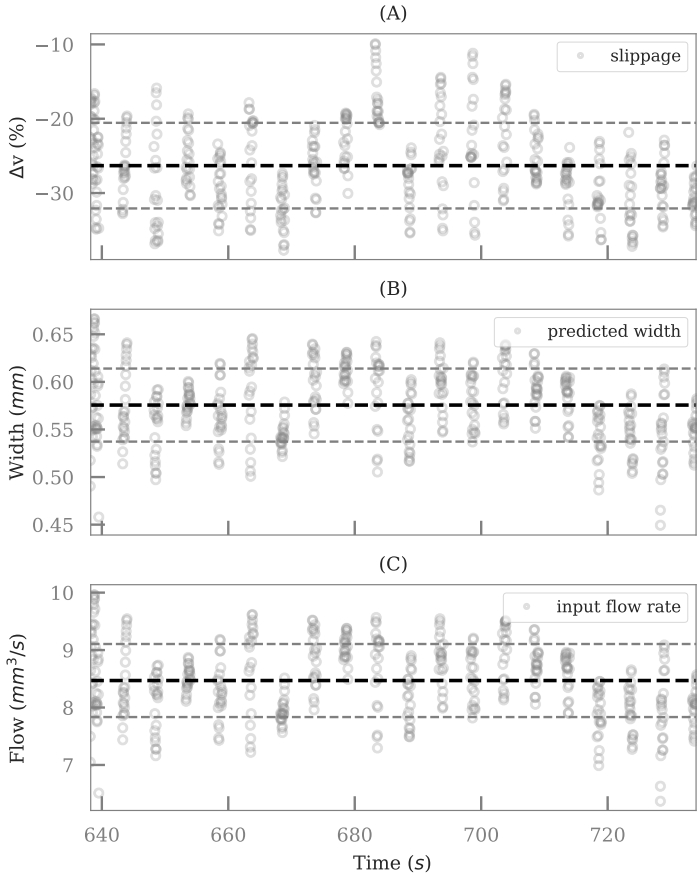


Figure 3.49: Derived values for a selected object in the print run, with (A) the feed speed slippage percentage, (B) the predicted track width and (C) the input flow rate.

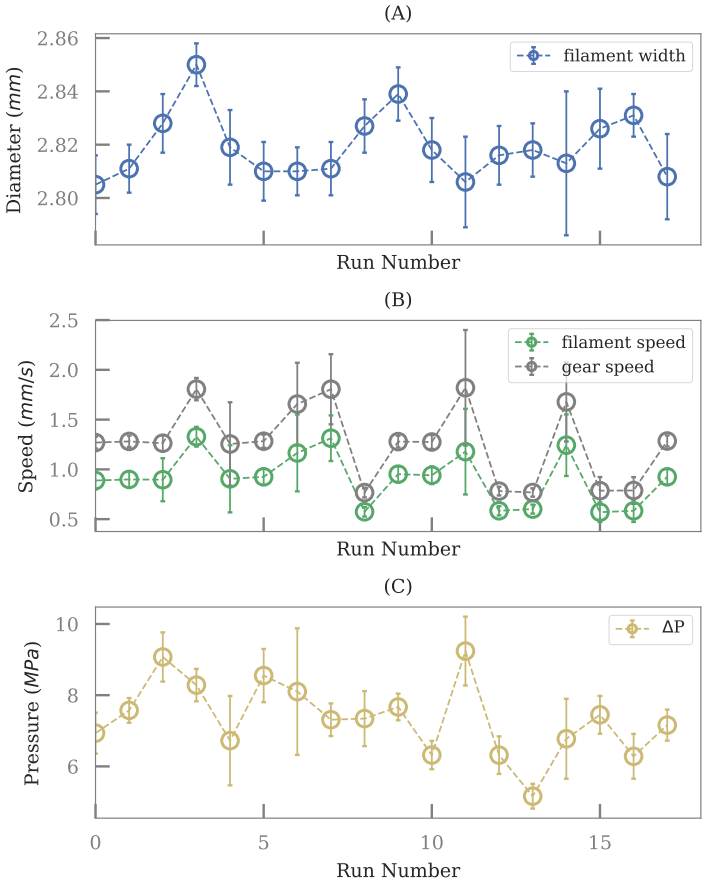


Figure 3.50: In-process monitoring results of all the test objects, with (A) the filament diameter, (B) the feed speed and (C) the pressure. Error bars indicate the standard deviation over all the interest lines per object.

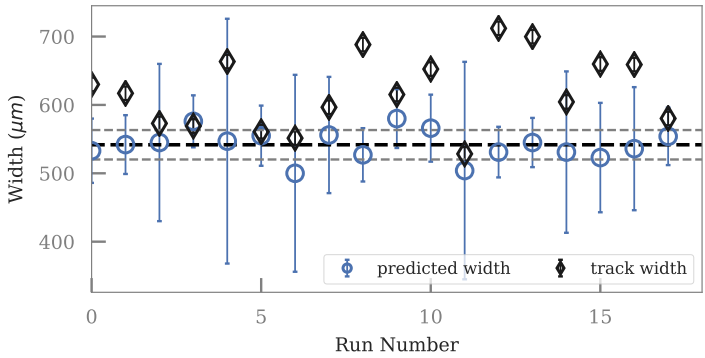


Figure 3.51: Comparison between the measured (unfilled diamonds) and predicted (unfilled circles) track width. The dashed horizontal lines indicate the mean and standard deviation of the predicted track width.

The two process input factors, applicable to the interest line in the SPO *Test B*, are the perimeter speed and the temperature. The in-processed measured process parameters are shown against these factors in Fig. 3.52.

Feed slippage, expressed as the difference between the filament and gear speed, shows a trend with respect to speed, but not with respect to temperature, except for the two outliers (indicated with diamonds at 200 °C and 215 °C). Note that the position of the outliers also represents the expected system response. That is, an increase in feed slippage at colder temperatures. It is, however, expected that the temperature effect will only become more significant at temperatures colder than 200 °C (see Section 3.1.9).

The feed slippage grouped by speed does show a clear trend, since the slippage increases with increasing print speed. The measured pressure drop grouped by the two test factors confirms the general model expectations of decreasing pressure at hotter temperatures due to the lower melt viscosity and the increasing pressure at faster speeds.

It is expected that the filament diameter variation is independent of both feed speed and temperature, as can be seen in the first box plot which compares the diameter with the temperature. The second plot indicates a possible trend, where the width seems to decrease with increasing speed. The estimated resolution limit (±10 µm) is also drawn on the plots with dash-dot lines and shows that this variation is less than the resolution limit. It is therefore assumed that the filament diameter variance is independent of the two input factors.

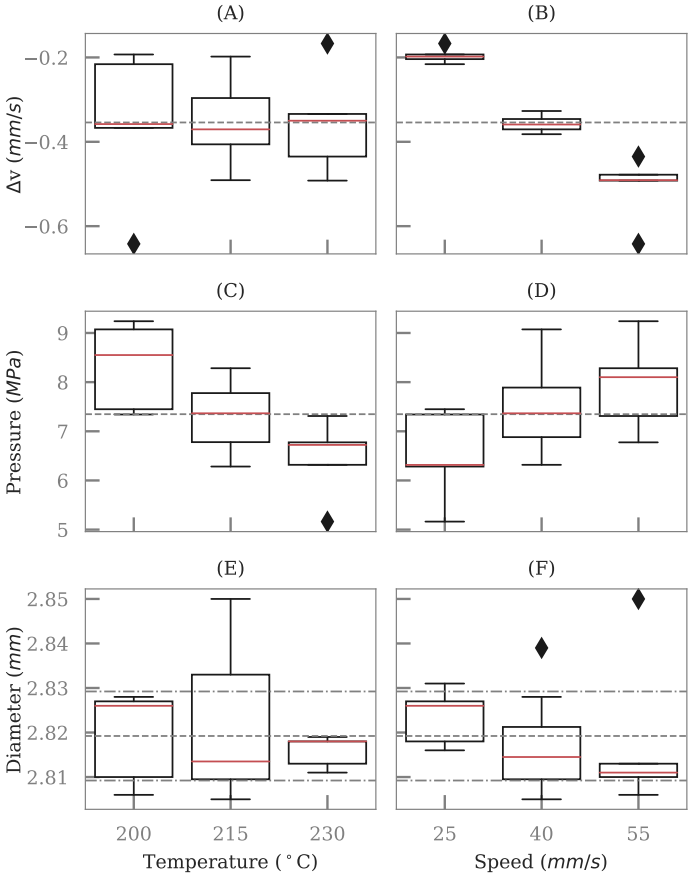


Figure 3.52: The three in-process measured parameters (rows: feed slippage, pressure and diameter) against the two input factors (columns: temperature and perimeter speed). Dashed horizontal lines are the mean for each parameter and diamonds outliers.

3.4.3 Liquefier Pressure Drop

The liquefier pressure drop is a key process variable. This variable is made even more interesting in that the extrusion force measurement sensor is non-obtrusive and already built into the RF1000 printer.

The pressure value, however, depends on many variables, including temperature, feed speed, liquefier geometry and material properties. Extrusion during printing is also not a steady state process, since most of the inputs are continuously varied, sometimes even with step changes.

The measured pressure, with respect to the commanded perimeter speed and temperature, is shown in the contour plot Fig. 3.53. Extrusion pressure is seen to be at its lowest in the lower right, at hotter temperatures and slower print speeds, while the maximum is in the upper left corner, at cooler temperatures and faster print speeds.

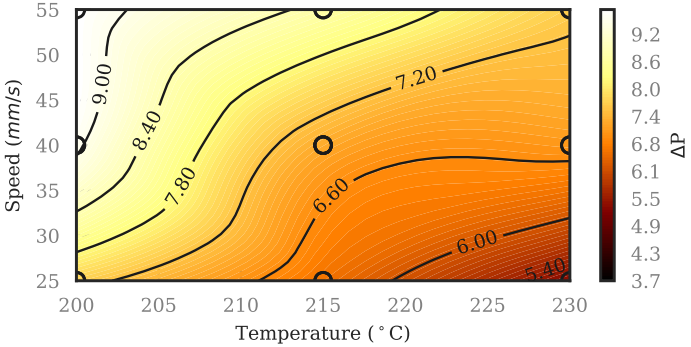


Figure 3.53: Measured pressure against commanded perimeter speed and extrusion temperature.

The liquefier pressure model of Bellini, Güçeri, and Bertoldi [102] (see Eq. (2.45)), in the form presented by [105, 138], was used to fit the material properties m, ϕ and α, while the geometric properties were assumed to be correct. The results are shown in Fig. 3.54, with circles indicating the mean measured pressure for the interest lines of each object and the solid lines indicate the model prediction for different speed levels.

Three material parameter sets were considered (*p0*, *p1* and *p2*), where *p0* is based on the values reported in [100] (see Table 2.8) and *p1* is the fit result. Fitting was done using the *SciPy* non-linear least squares optimisation method *curve_fit* [178], with *p0* as the initial values. Both the *p1* and *p2* parameter sets seem to approximate the measure values better than the *p0* set.

The third parameter set (*p2*) modifies the Bellini equation by making the length of the first liquefier section dependent on the melt front location (see Section 2.8.11).

The melt front depth (z'_m, the distance of the melt front location from the filament entry point) was subtracted from the total liquefier length to obtain $L1 = f(L_L, T, v_f)$, with Eq. (2.55), where L_L is the liquefier length from the entry point to the conical section start, T is the liquefier wall temperature and v_f is the filament entry speed. The thermal diffusivity (α_d) of PLA was taken as $8.8245 \times 10^{-8}\,\mathrm{m}^2/\mathrm{s}$ and the liquefier length (L_L) as 22 mm.

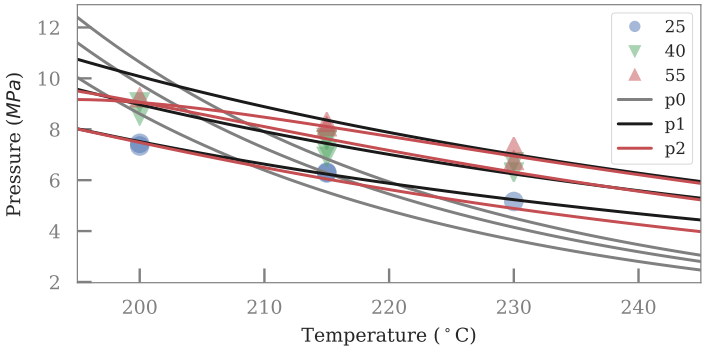


Figure 3.54: Liquefier pressure model fitting results, with unfilled circles indicating the measured pressure at the commanded perimeter speeds (in mm/s), and model values with parameters sets *p0*, *p1* and *p2* at 0.6 mm/s, 0.9 mm/s and 1.3 mm/s input feed speeds.

The temperature dependent gradient for the *p0* parameter set is too high. The fit optimisation therefore decreases material activation energy (α). The values of each set are given in Table 3.14, where $L_1(L_L, T, v_f)$ indicates whether the $L1$ melt front adjustment was used. Parameter set three (*p2*) further reduces both m and ϕ, but increases α.

The effect of these material parameters are illustrated in Fig. 3.55, where the total pressure drop, as well as the pressure for each segment, are shown as predicted by the model with the different parameter sets. Parameter set *p1* reduces the pressure temperature gradient as compared to *p0*, while the biggest resistance to extrusion is predicted to be in the conical nozzle section and the lowest at the liquefier entry section.

The third parameter set finds a different solution, with the largest pressure drop across the entry section, since the length of this section was increased considerably. Note that the entry section length, filament diameter and the material thermal diffusivity were assumed to be constant and correct.

The fit by the third set does not seem to be correct, according to expectation that the pressure is inversely proportional to the nozzle radius with R_n^4. Nevertheless, the entry length reduction apparently models the

Table 3.14: Parameter sets used for the pressure drop model.

x_i	**p0**	**p1**	**p2**	**unit**
m	4.000	2.888	1.270	
ϕ	3.33E-19	1.13E-13	1.53E-03	
α	6805	2873	3898	K
L_1		22.0		mm
L_3		1.0		mm
R_1		1.5		mm
R_3		0.3		mm
β		60.0		degree
$L_1(...)$	False	False	True	

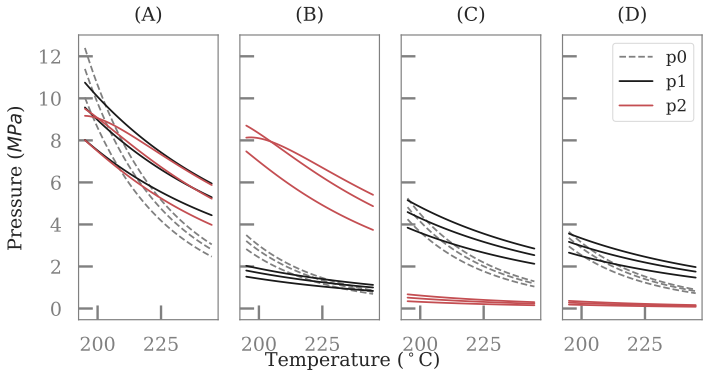


Figure 3.55: Pressure drop for each liquefier segment for the three model parameter sets (*p0*, *p1* and *p2*), with (A) the total pressure drop, and (B) the liquefier entry section, (C) the conical section and (D) the nozzle tip section.

more extreme case of lower temperature and higher speed better (see upper left area in Fig. 3.54) and attempts to model the fact that the melt length is not constant.

The modelled melt length is shown in Fig. 3.56, as a function of temperature and feed speed. It increases with temperature and decreases with speed, as expected. This means that the pressure drop will *increase* with increased temperature or decreased speed (contrary to the standard principle), since the length of molten material will be longer.

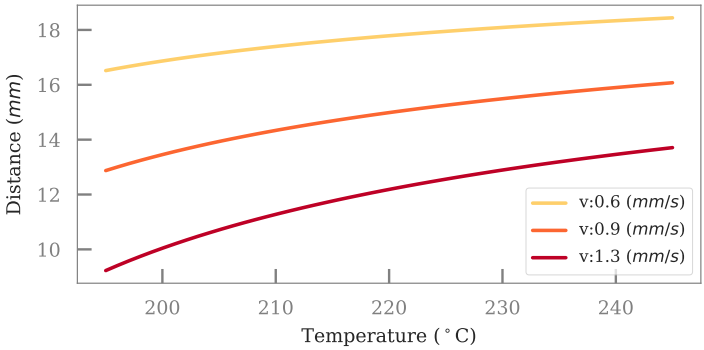


Figure 3.56: Liquefier melt length ($L1$) variation due to the melt front position dependency on temperature and feed speed, for parameter set *p2*.

The residuals of the three parameter sets are shown in Fig. 3.57 and both *p1* and *p2* achieve a lower residual error, which is balanced around zero. The constant melt length parameter set (*p1*) is investigated in more detail next, since it not clear if residual error improvement gained by adding a fit variable in *p2* models the actual process better.

Four plots of the *p1* parameter set residuals provide further details on the fit result in Fig. 3.58.

The distribution of the residuals appear to be normal and centred around zero, while the lag and the run plots indicate that there are no significant temporal drift effects, which have not been captured by the model and the fitted parameters.

3.4.4 Feed Slippage Prediction

Closed loop control of the feed mechanism requires the real time measurement of the feed slippage. This can be done with an optical method, as presented in Section 3.1, or with a rotatory encoder.

Mounting the optical system can, however, be difficult due to space constraints. An encoder system on the other hand can be made smaller, but requires at least two encoders for both the gear and filament speed.

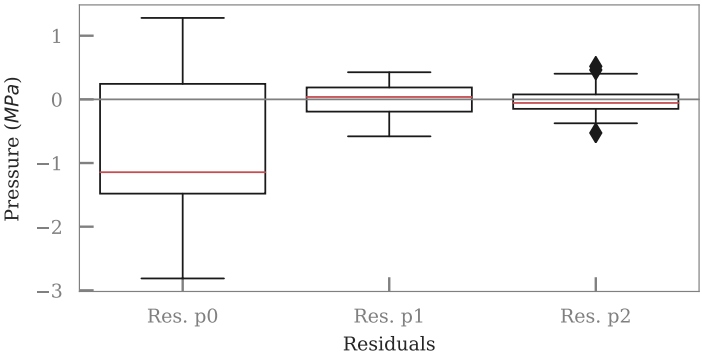


Figure 3.57: Comparison of the liquefier pressure fit residuals for parameters sets *p0*, *p1* and *p2*.

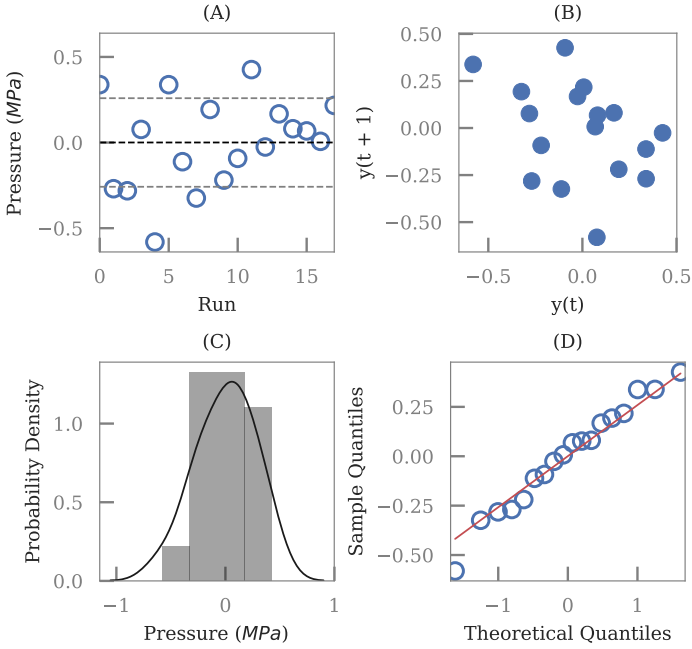


Figure 3.58: Four plots of the liquefier pressure model fit residuals for parameter set *p1*, with the (A) run sequence, (B) lag, (C) histogram and (D) Q-Q plot.

Furthermore, fast high resolution encoders are also expensive. An alternative option is to predict the feed slippage using non-direct measures. For example, the liquefier pressure or accelerometer readings can be used (see Section 2.9).

The RF1000 is conveniently fitted with a pressure sensing ability and the concept of using it to predict the feed slippage is investigated next.

The Bellini model can be rewritten so that the input speed v_f can be predicted, as shown in Eq. (3.14). The constants (A_1, A_2 and A_3), as defined by comparing Eqs. (2.45) and (2.46) with Eq. (3.14), were found to be 514.893, 1306.326 and 903.091 (rounded to three decimals) for the parameter set *p1*.

$$v_f = \phi \left(\frac{\Delta P}{(A_1 + A_2 + A_3)\, H\,(T)} \right)^m \tag{3.14}$$

The estimated input speed with Eq. (3.14) is only a function of the temperature and pressure, along with the model constants. It is therefore possible to predict the current feed slippage, by comparing the calculated input speed with the commanded feed speed, without additional measurement systems.

The feed speed difference was also modelled empirically, with the measured data shown previously in Fig. 3.50, with Eq. (3.15), where β_0, β_1, β_2 and β_{12} were determined using the *SciPy* non-linear least squares optimisation method *curve_fit* [178]. Initial values for the parameters were all set to one.

$$v_f = \beta_0 + \beta_1 T + \beta_2 \Delta P + \beta_{12} \Delta P T \tag{3.15}$$

The coefficients for Eq. (3.15) found with the fit routine are given in Table 3.15 (note that the reported values are rounded).

Table 3.15: Coefficients for the empirical model for v_f.

$\beta_\mathbf{i}$	**Value**	**Unit**
β_0	9.99×10^{-3}	mm/s
β_1	-1.28×10^{-1}	mm/(s · K)
β_2	2.03×10^{-3}	mm/(s · Pa)
β_3	-3.49	mm/(s · K · Pa)

The free air feed slippage data measured in Section 3.1.9 were also modelled with a similar method to represent the feed slippage in millimetres per second. The measured speed, as well as the filament gear

speed difference, are compared against the predicted values of all three models in Fig. 3.59. The speed difference was calculated by subtracting the measured gear speed. This can be replaced with the commanded feed speed.

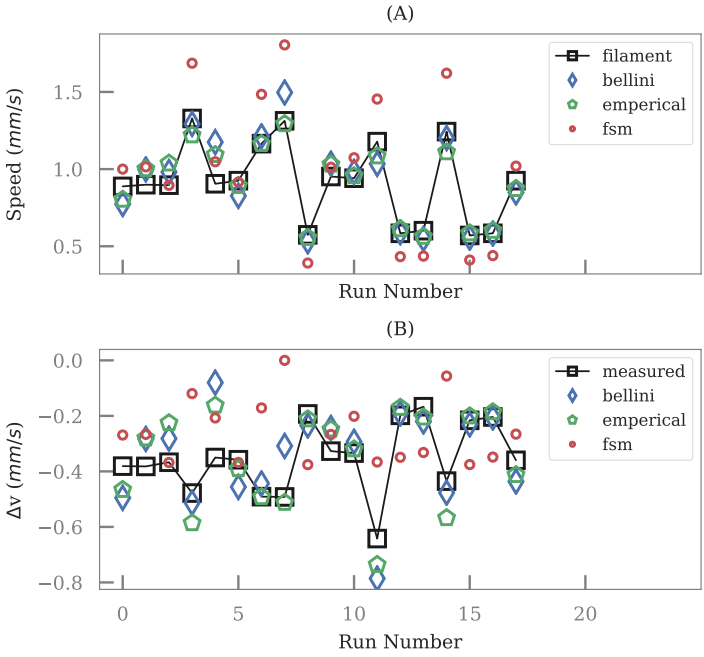


Figure 3.59: Comparison between the measured speed (unfilled squares and solid line) with the reversed Bellini model (unfilled diamonds, with parameter set *p1*), the fitted empirical model (unfilled pentagons) and the empirical feed slippage model (unfilled circles). Plot (A) is the measured speed and (B) the speed difference from the measured gear speed.

Boxplots of the residuals for the predicted speed of each model are shown in Fig. 3.60. Both the empirical and the Bellini fits achieve a near zero median, while the empirical feed slippage model (*fsm*) has a larger variation and a non-zero median. Note that the *fsm* tests were done in free air, while the other two models are based on the average interest line variables. The *fsm* measurements were also performed at faster speeds than the print tests.

Only empirical and reversed Bellini models are investigated further, since they show smaller zero centred residuals. Contour plots of the data are shown in Fig. 3.61, which provides a three-dimensional perspective of the feed slippage characteristics.

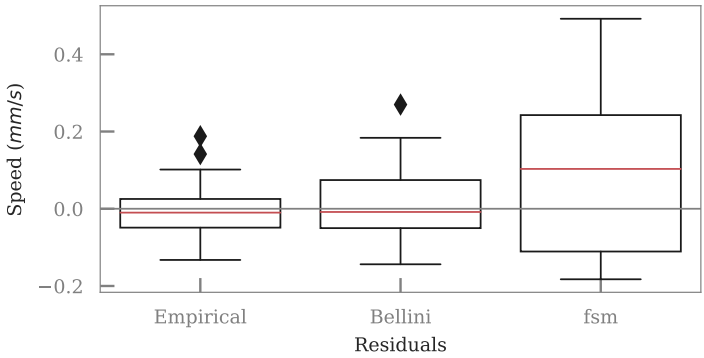


Figure 3.60: Comparison between the residuals of the input feed speed models, where *fsm* is the feed slippage model.

The maximum feed slippage is in the upper left corner (cooler temperature and faster print speed) and the minimum in the lower right hand side of each plot. The contour lines are approximately horizontal, as already indicated by the boxplots of speed slippage in Fig. 3.52. Both the empirical and the analytical model predict parabolic contour lines, with slightly higher slippage in the middle at around 215 °C, which is also an unexpected result.

3.4.5 Model Validation

The empirical and analytic models were validated by comparing the predictions of the models with the measured values, for *Test V* and *Test R*, where *Test V* is the verification test performed in Section 3.3.14 at the selected operating point of 55 mm/s and 230 °C for the perimeter speed and extrusion temperature. *Test R* is the reverse test performed in Section 3.3.12 at the centre run settings of 40 mm/s and 215 °C for the perimeter speed and extrusion temperature.

The three measurement data streams were combined for each run, the interest lines inferred by the fitted commanded speed values and the average data of all the interest lines segment of each object were retrieved. The resulting data sets, which describe each object, were then combined in a single file.

The result for the width measurements of the verification runs are compared to that of the DOE run (*Test B*) in Fig. 3.62, with dashed lines indicating the mean width of the two experiment sets (training and verify data).

The difference between the two is approximately equal to the estimated width resolution limit of ±10 µm, while the total variation is within the manufactured specified limits, albeit lower than the centre value of

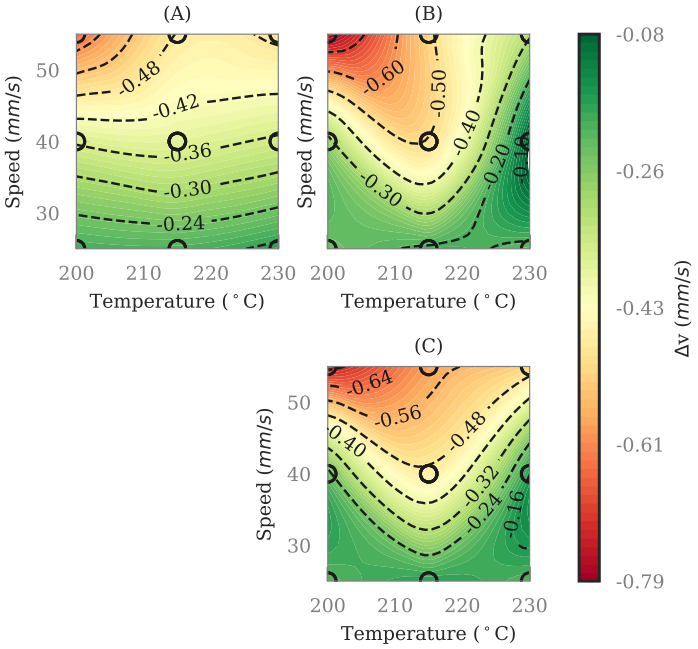


Figure 3.61: Contour plot of the speed deviation for the (A) measured values, (B) the reversed Bellini model with parameter set *p1* and (C) the empirical model, against the commanded perimeter speed and liquefier temperature.

2.85 mm (see Section 3.1.7). It is therefore assumed that the main input into the system, the material, is in control and within normal variance, which allows for the other in-process measured parameters to be compared.

The measured pressure for the verification runs are shown in Fig. 3.63, along with the predicted pressure with parameter sets *p1* and *p2*, given the measured input feed speed and set point temperature.

The first five experiments are that of *Test V*, at a faster speed and a higher temperature, than the rest of the experiments (*Test R*). Both model parameter sets predict a higher pressure for these five experiments, except for the third object.

Note that the expectation is that all five experiments should have the same pressure, since the input factors are the same. But this is not the case, as seen in Fig. 3.63. Both models also predict an approximately equal pressure for the first five tests.

The remaining data points were measured in the reverse print run,

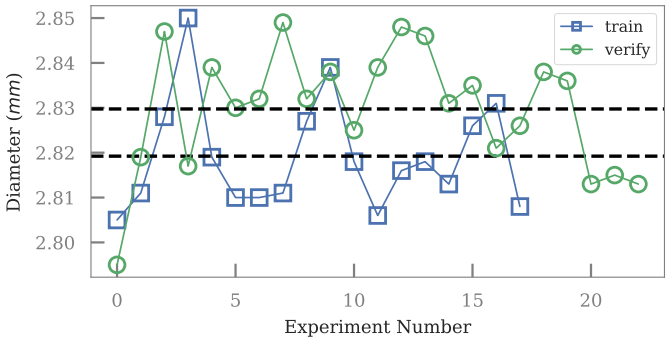


Figure 3.62: Measured filament diameter comparison between the verification and training data sets. The lower horizontal dashed line is the mean train data width and higher line is the verification mean width.

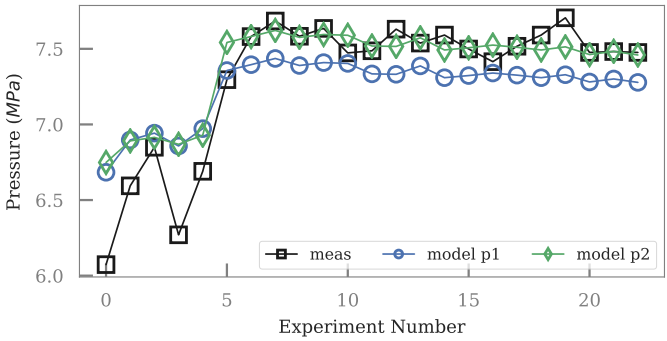


Figure 3.63: The measured pressure for the verification runs (unfilled squares) compared with the fitter model predicted pressures with parameter set *p1* (unfilled circles) and *p2* (unfilled diamonds).

which were performed at a lower temperature and a slower speed. Here, the parameter set *p1* predicts a lower pressure, while the *p2* set seems to perform better. Note that the measured pressure shows less variance over all the reverse run experiments.

Possible reasons for the larger variance of the first five experiments are an increase in less predictable slippage behaviour at higher speeds and unidentified (or neglected) input factors such as the material water content, filament diameter, dirt build up in the nozzle due to burned material or the extrusion history. *Test R* was however performed after *Test V*.

The input speed, given the temperature and pressure, was also modelled with the *p1* parameter set and is shown in Fig. 3.64. Both the input and the speed deviation are plotted and the input speed is seen to vary more for the first five experiments. The modelled speed also shows more variance than the measured speed. It does, however, give a reasonable prediction of the speed.

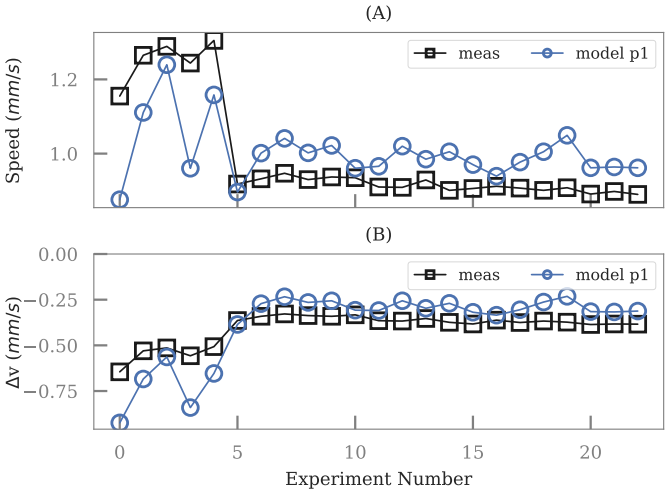


Figure 3.64: Verification test data of (A) the input (filament) speed and (B) the feed slippage (filament minus gear speed), where the measured values are unfilled squares and the predicted values unfilled circles.

Finally, the residuals for the verification prints are shown in Fig. 3.65 for the pressure and input speed models. All the residuals approach zero, but only the parameter set *p2* achieves an approximately zero median. Furthermore, outliers are also indicated with the diamonds for each model.

The pressure measurement was estimated as ±0.26 MPa and the speed measurement uncertainty was ±0.06 mm/s, for a 2 mm/s extrusion (see Sections 3.1.8 and 3.2.4). The quantile boxes shown in Fig. 3.65 agrees well

with these uncertainty values, which motivates that the model parameters are working for such uncertainties and factor ranges.

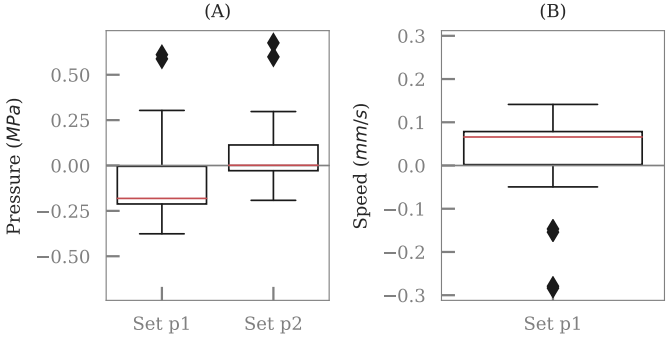


Figure 3.65: Verification prints in-process data models residuals for (A) the pressure and (B) the input speed, with parameter sets *p1* and *p2*.

3.4.6 Implementation and Application Options

The models with the fitted parameters can be used to derive a reference table for static adjustment of either the slicer or the firmware execution of a **G1** G-code command. In-process results can also be used to detect wear and inform maintenance plans, while the methods presented here, especially with the Single Print Optimisation (SPO) technique, can be used to characterise new extruder designs or filament material. Furthermore, in-process measurements can also be used for quality monitoring and quality assurance of the final part.

Another interesting application is closed loop control of the feed mechanism without the need for any additional sensors on the RF1000 or similar printers. This should be possible by using the input feed speed model Eq. (3.14). The feed multiplier approach can be used to modify feed speed as was done by the machine vision print server interface in Section 3.1.10. Alternatively, this model can be linearised over a smaller temperature range and implemented in firmware to control the feed speed, using a standard PID feed-back loop. This loop will vary the speed of the feed mechanism, and correspondingly that of the Bresenham driving axis (see Section 2.6.4) to ensure constant volumetric flow, while the total travel distance is achieved independently, by only counting the number of steps regardless of the speed.

The speed variation will, however, require the recalculation of the number of steps required for deceleration toward the trajectory end speed. This costly calculation can be avoided using the *time per step* instead of the *velocity per step* approach (see Section 2.6.3).

4 Vapour Deposition Fabrication

The concepts learned from Fused Filament Fabrication (FFF) are used to develop a Vapour Deposition Fabrication (VDF) printer in this chapter. The aim of a VDF printer is to produce nanometre to micron sized, complex, three-dimensional structures, which can be, for example, new sensors. Another application area is directly printing electrical connections on a substrate, without needing any lithographic processes.

A short overview of Physical Vapour Deposition (PVD) is given, with particular attention to dynamic stencil lithography. This is followed by defining the requirements and the main components for a VDF printer, which includes the design of the material source. The last section reports the first test results of such a printer.

4.1 Dynamic Stencil Lithography with Physical Vapour Deposition

Physical Vapour Deposition (PVD) is conventionally used for the application of a thin, homogeneous coating of material on a surface. The possibility of creating inhomogeneous structures with dynamic stencil lithography using PVD, though, are also of interest. This is also the VDF base process. This section briefly describes PVD with thermal evaporation and the three main limitations of dynamic stencil lithography.

A cost effective PVD technique is the evaporation of the source material with a resistive heat source, as shown in Fig. 4.1. The source material is evaporated by either directly or indirectly heating it. This vapour travels in straight lines to the substrate, where it condensates, which creates a thin, homogeneous, film on the substrate when several factors are carefully controlled.

An important factor is the evaporation rate, which depends on the temperature and the vacuum pressure level. The heat is generated by a high current flowing through the resistive heating material, which can

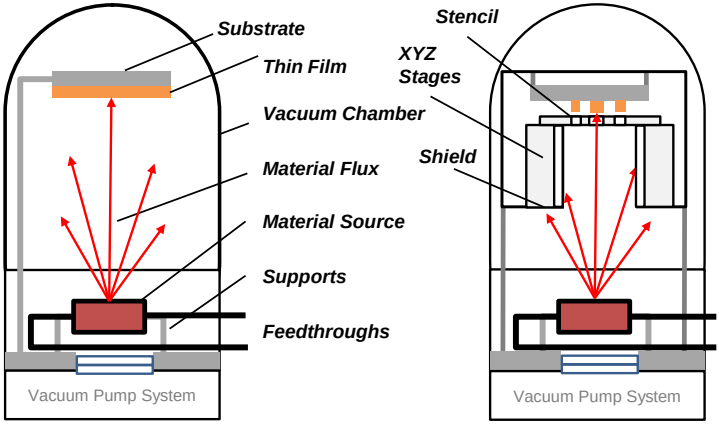


Figure 4.1: Physical Vapour Deposition: (left) main components and (right) the concept of quasi-dynamic stencil lithography (based on [188]).

be the source material itself, a material boat or a heater wire wrapped around a crucible, for example.

The material source and the substrate are placed in a vacuum chamber, which is traditionally a bell like jar, as shown in the figure. A vacuum environment is essential to achieve the deposition of thin films, since it ensures that the evaporated particles travel in straight lines from the source to the substrate and create a uniform layer. This is due to the increased mean free path of the particles in the vacuum, which is the average distance travelled between successive collisions with residual gas molecules. The vacuum also reduces the risk of deposition contamination and alloying.

The mean free path (λ) can be estimated with Eq. (4.1), where d_m is the diameter of the molecule, T is the temperature, P is the pressure and k_B is the Boltzmann constant.

$$\lambda = \frac{k_B T}{\sqrt{2}\pi P d_m^2} \tag{4.1}$$

The evaporated particles will travel straight lines until they reach a surface, if the vapour is in the molecular flow regime. The flow regime is estimated with the Knudsen number (λ/d), where a number larger than one indicates molecular flow and d is the orifice diameter. This is because inter-molecular collisions are rare and collisions with orifice wall dominate [189].

4.1.1 Evaporation Rate

The vacuum environment reduces the heating power required to evaporate the material, since convection heat losses are negligible in vacuum. The required power is further reduced, since the material is only heated to a point where the material vapour pressure is larger than the residual gas pressure.

The equilibrium vapour pressure, also called the saturation pressure, is the point where the vapour pressure of the material is in equilibrium with the liquid surface, in a closed container. That is, the number of atoms leaving the surface is equal to the number that are returning, at a given temperature [190].

The rate of evaporation, therefore, depends on the residual gas pressure level and on the temperature, where a small change in temperature around the critical evaporation point can make a large difference in the rate, since the vapour pressure is strongly temperature dependent [190].

The total evaporation flux (φ) can be modelled by the Hertz-Knudsen equation, given in Eq. (4.2).

$$\varphi \equiv \frac{dN}{Adt} = \alpha_v \frac{\left(P^* - P\right)}{\sqrt{2\pi m k_B T}} \tag{4.2}$$

The flux is defined as the rate of the number of evaporating atoms per time (dN/dt), per orifice cross section area (A). This can be used to approximate the free surface evaporation rate of a thermal source in PVD, where m is the mass of the vaporised material, k_B is the Boltzmann constant and T is the absolute temperature. The pressure differential ($(P^* - P)$) is the difference between the material vapour pressure (P^*), at temperature T, and the pressure of the vapour above the surface (P) [190, 191].

This equation models the dependence of the evaporation rate on the source temperature (T) and the vacuum chamber pressure (P), where the maximum rate occurs at $P = 0$ and $\alpha_v = 1$. The correction factor (α_v), also called the sticking factor, is determined experimentally [192].

The actual evaporation rate, however, will be one-third to one-tenth of the calculated rate, due to collisions above the evaporating surface and surface contamination, amongst other effects [190].

4.1.2 Inhomogeneous Thin Films

The fabrication of thin films generally aims to produce a very homogeneous layer on the substrate. The substrate is therefore placed relatively far away from the source, to ensure a uniform coating. The on purpose generation of inhomogeneous depositions, however, is also of interest.

This can be achieved by exploiting atomic shadowing, an effect which is avoided in thin film deposition. Atomic shadowing occurs when the substrate is titled with a sufficient angle (>70°) from the material flow direction. Glancing angle deposition (GLAD) and substrate motion were,

for example, used by Robbie and Brett [193] to create "nano-structurally controlled sculptured thin films".

GLAD, alternatively called Dynamic Shadowing Growth (DSG), can be used to create overhang features or helical structures [194]. An important factor to achieve such structures is that the self-shadowing effect becomes the dominant growth mechanism. This requires a highly directional material vapour beam, which can be produced with thermal evaporation, electron beam evaporation or molecular beam epitaxy and pulsed laser deposition (PLD), for example [194].

Shadowing can be used to create even more localised deposition if a stencil or mask is placed between the source and the substrate. The first stencil techniques were used as early as 1969, where the stencil was clamped to the substrate [195]. This was significantly extended by Savu *et al.* [188], who moved the stencil relative to the substrate with a motion stage, to achieve dynamic stencil lithography (see Fig. 4.1).

4.1.3 Dynamic Stencil Lithography

Stencil lithography allows resistless nano-fabrication and the review by Vazquez-Mena *et al.* [195] finds that this has many advantages, including that no resist processing is required and that easy implementation is possible. The stencil can also be re-used and dynamic stencil lithography is also possible. Applications listed in [195] include transistor fabrication on non-planar substrates and magnetic nano-structures fabrication.

Stencil lithography has been realised with different material sources, including thermal evaporation, PLD [196] and molecular beam epitaxy [197].

The main limitations of stencil lithography are, however, clogging, blurring and stencil stability [195]. Both clogging and blurring are schematically shown in Fig. 4.2. The stencil itself must also be mechanically, chemically and thermally stable.

Stencil Clogging

The vapour will condense on the stencil, which can close the aperture and reduce the deposition resolution during fabrication. This can be advantageous and it can be used to create small features, but it is in general not desired. Clogging becomes significant if the deposition thickness is comparable or larger than the aperture size [195]. This also reduces the operational life span of the stencil.

Clogging can be reduced by heating the stencil [198]. A special stencil coating can also be applied to reduce clogging, while cleaning the stencil can increase the operational lifespan [195].

Deposition Blurring

Blurring of the deposited material is probably the biggest challenge in terms of achieving high resolution [198]. This is caused by the material vapour beam divergence after passing through the stencil aperture and is

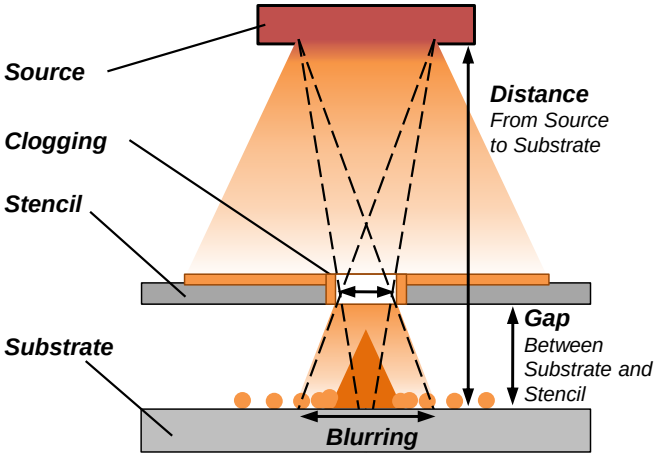


Figure 4.2: Stencil clogging and deposition blurring.

proportional to the gap between the stencil and the substrate (G), with $B \propto G(S/D)$, where B is the blurring amount, S the source size and D the distance between the source and the substrate [188].

The material will also diffuse on the substrate surface, which will further increase the amount of blurring. The size of blurring can even be bigger than the aperture size in the nano-scale regime. Additionally, the amount of blurring depends on the source material, deposition rate and the substrate temperature [195].

The first step to address blurring is to reduce the clearance between the substrate and the stencil. This is, however, limited by the curvature of both and the parallelism between them. Blurring can also be reduced by the careful selection of the substrate material, coating of the stencil and by using a mobility-inhibiting gas [195].

Deposition Rate

The stencil developed by Savu, Xie, and Brugger [198], uses a resistive heater to reduce clogging and to bend the aperture towards the substrate. In their experiments they deposited aluminium at an evaporation rate of around 0.2 nm/s, with apertures ranging from 100 – 200 nm.

The evaporation rate not only limits the size of structures, but also sets temporal stability requirements for the process. For example, the motion stage was found to have a drift of 2.2 nm/s.

Stencil Considerations

In general, the aim is to obtain nanometre resolution structures with dynamic stencil lithography. This requires that the aperture size is as small as possible. This, however, reduces stencil thickness and therefore the mechanical stability. Reducing the aperture size also increases the clogging effect.

The gap between the substrate and the stencil should also be as small as possible. This requires precise alignment, positioning and stability of both the substrate and stencil.

4.2 The Vapour Deposition Fabrication Concept

A special version of dynamic stencil lithography with thermal evaporation is shown in Fig. 4.3. Here, a very localised and miniaturised thermal evaporation source is mounted in close proximity to the stencil and the substrate. Note that this method is differentiated from the process described by Wallenberger [199], since this method uses a thermal evaporation source and not the more complex laser-assisted chemical vapor deposition (LCVD) process.

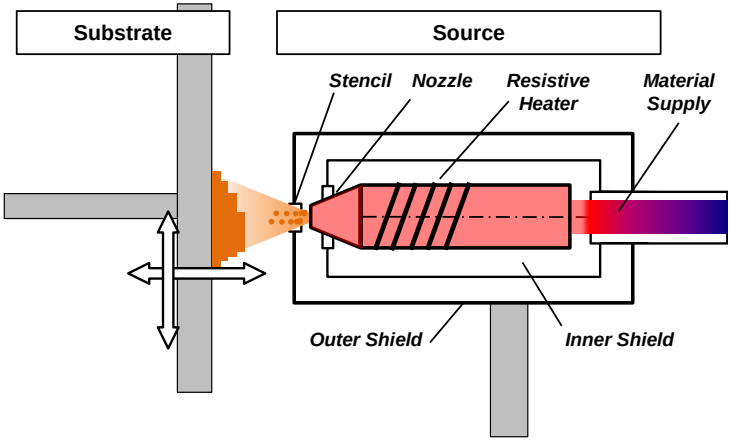


Figure 4.3: Three-dimensional micro-scale additive manufacturing with thermal Vapour Deposition Fabrication.

The material is supplied, similar to FFF, with a wire, which is fed into the evaporator. This can be replaced with a granulate feeding mechanism, especially if the source material heat conduction is too high.

The evaporator has a conical section with a small exit diameter and a heater wire is wrapped around it. A method to feed the heater wire into the inner shield is also required.

The source exit diameter can be further reduced, if a stencil is placed between the source and the substrate.

Thermal radiation from the source is reflected back with the inner and outer shields. This is important to reduce heat losses and to keep the source compact.

The source itself is fixed to the system frame, and not to the motion stages, to increase the thermal conduction path length to the temperature sensitive positioning stages. The substrate is therefore moved in the X, Y and Z-axes.

The concept works in principle as follows. A thermally isolated feed mechanism feeds material into the evaporator, which conducts the thermal heat generated by the heater wire to the source material. This material is then evaporated by heating. The vapour travels in a straight line towards the source, due to the vacuum environment. The vapour is focused on the substrate with the nozzle tip, while the substrate is moved in the X, Y and Z-axes.

It is also possible to add a heated source, a substrate gas supply (for example oxygen) and a heated stencil to the concept presented in Section 4.1. Furthermore, active cooling can be achieved using cold water pipes connected to heat sensitive components. A high resolution ($\pm 10\,\text{nm}$) displacement sensor between the substrate and the source will also be useful as well as methods to monitor the build progress. These are only noted here, since the focus of this work is to show a proof of concept.

4.3 Micro-Printer System Design

The realisation of the VDF printer concept, called the micro-printer in this work, integrates several sub-systems. These must interface with each other and work together to achieve the main system requirements. This section therefore defines the system requirements, which are used to select or to design the main components. These requirements are also used in Section 4.4, where the design of the material source is presented.

4.3.1 General System Requirements

Requirements for a VDF system are given in Table 4.1 and were derived based on the stencil lithography process limitations (clogging, blurring and stencil stability) and on the VDF concept.

The high vacuum environment is a trade off between what is reasonable to achieve and reducing the mean free path length, while the required heating power is a compromise between heat loss and the evaporation rate.

If there was no heat loss from the source to the rest of the printer, then an infinitely small heater could be used. The heater, however, needs to be connected to the rest of the printer and some heat loss through conduction and thermal radiation is inevitable. The approximate numeric values given in Table 4.1 are based on the modelling results found in Section 4.4.

Table 4.1: General system requirements for the micro-printer.

Requirement	Specification
High vacuum environment	Around 10^{-6} mbar
Sufficient thermal power	Must be more than the thermal losses, not too much to reduce heating of sensitive components and enough for a reasonable evaporation rate, in the range of 10 – 50 W
Reduce heat loss	Very localised, concentrated heating, with a maximum conduction loss of 10 W
High resolution	Nozzle diameter and positioning in the range of 0.1 – 50 µm
Reduce blurring	Stencil substrate gap must be less than the nozzle diameter
Reduce clogging	Include the possibility to heat the stencil
Temporal stability	Reduce drift over time to less than 10 % of the resolution
Cost effective	Re-purpose an existing vacuum chamber and other parts already in the laboratory, and a reusable source
Standardisation	Use concepts of FFF, including the firmware

The main differentiating factor between thermal PVD sources and the VDF source is that the material source is small and should use as little as possible power to evaporate the material. This requirement is emphasised by the maximum operating temperature of the positioning stages of 60 °C. The use of currents, for example of less than 5 A, also reduces the safety risk, as compare to the conventional 50 – 400 A rating of PVD thermal heaters. This also reduces the cost of this system.

A smaller thermal power input into the system should also reduce the effect of thermal expansion on the alignment. *Temporal stability* requirement implies that the material source should not drift relative to the substrate. This is essential, since (1) the expected deposition rate is low (see Section 4.1.3), (2) high resolution is required and (3) the current system uses open loop control.

The type of materials used to construct the system is important. The following points should be considered during material selection:

i Operation temperature (800 – 1400 °C)

ii High vacuum environment compatibility (low out-gassing and low vapour pressure)

iii Chemical reaction – it should not contaminate the source material (e.g. alloy with it)

iv Standard profile availability or machining options

It is possible to use more exotic materials in future work, with for example very low thermal expansion, to improve the temporal stability of the system. Such materials, though, tend to be more expensive and harder to machine.

Any parts close to the thermal evaporation source should not evaporate and create unwanted deposits. Neither on the heating system nor on the substrate. Furthermore, all the components inside the vacuum chamber should not create *virtual leaks* by out-gassing.

Stainless steel, for example, can be used up to 870 °C, before it starts to oxidise in atmosphere, while it will melt around 1400 °C. It is also recommended that the free machining stainless steel grades (303, 303S, 303Se, 416) are only used for small parts, such as fasteners, due to the high sulphur or selenium content [200]. Care should also be taken with the selection of fastener materials. Brass, for example, is commonly used since it is not expensive and easy to machine, but it can have a high zinc content. Zinc has a high vapour pressure and can evaporate, even at moderate temperatures. Nickel plating can reduce the zinc out-gassing [200].

4.3.2 Main Components

The presented requirements and limitations were used to derive main components for the printer, as shown in Fig. 4.4. Three component categories can be discerned from this diagram namely: the electronics, vacuum chamber and the vacuum pump system.

The electronics interface with the machine and user commands, while it also controls the vacuum system, positioning stages and the evaporation source. An open source printer controller board was selected to control the printer (see Section 4.3.5), while a purpose built controller is used for the vacuum system (see Section 4.3.3).

This design is also influenced by the *cost effective* requirement. The vacuum chamber, for example, as well as the vacuum pumps and gauges were already available and could be redeployed in this system. This reduces cost and time needed to build a new chamber.

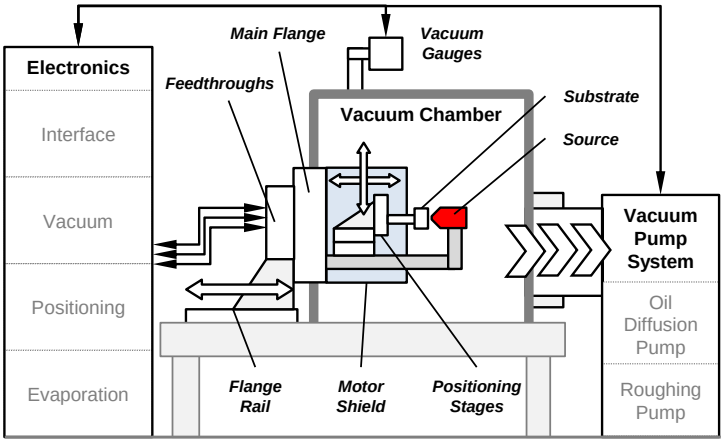


Figure 4.4: Main components of the VDF micro-printer.

The Main Flange

The vacuum chamber has several flanges, which are used to mount the pressure gauges and to connect the vacuum system. A flange, with an external diameter of 225 mm, is used to mount the base plate, which supports the whole printer, as shown in Fig. 4.5. This flange is called the *main flange*.

The three positioning stages, which actuate the substrate and the source are mounted on the base plate. A thermal radiation shield (not shown) is fitted around the stages. This protects the stages and the electronics from stray vapour and thermal radiation. It also protects the cables when inserting the printer into the chamber.

The main flange is mounted on the flange rail, which is used to insert the printer into the chamber. This allows for easy access to the printer when the flange is positioned outside the chamber as well as improved reproducibility of the printer location inside the chamber. This is important if it is required to interface the main flange with other flanges, for example, to align a laser beam or an optical monitoring system.

Mounting the whole printer on one flange was motivated by the *temporal stability* and the *high resolution* requirements, since this will enable fine alignment of the substrate and the source, before it is inserted into the chamber. Furthermore, the effect of vibrations on the relative distance between the source and the substrate will be reduced.

Using a single flange, however, constrains the available area for mounting all the feedthroughs, the shield and the base plate. Careful design is therefore required to ensure that all the components fit on the flange. The feedthrough connector details are given in Table 4.2. Only the nitrogen, which is used to vent the chamber, is supplied with a separate flange.

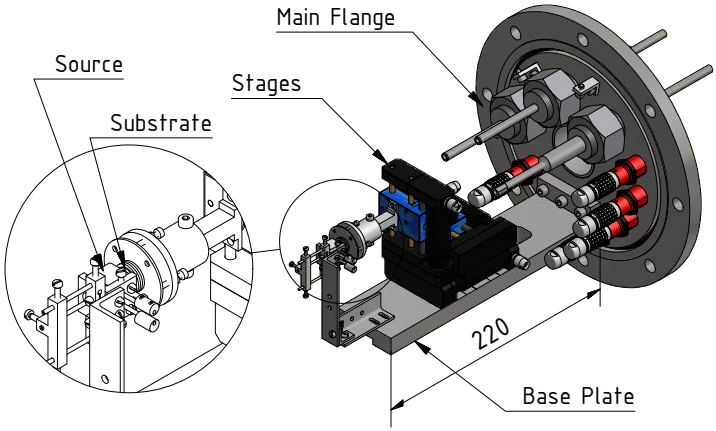


Figure 4.5: CAD of the micro-printer main flange, with the base plate supporting the three positioning axes, which position the substrate mount. Detail view (left) shows the source and the substrate.

Table 4.2: Vacuum camber feedthroughs

Function	**Connector**	**Number**
Motion Stages	15 pin Lemo® plug	3
Peripherals	15 pin Lemo® plug	2
Heater: Source	High current	1
Heater: Substrate	High current	1
Oxygen	Swagelok®	1
Cooling Water	Swagelok®	2
Nitrogen	Separate flange	1

Note that the cooling water requires two connections (inlet and outlet). Mounting space for the five 15-pin Lemo® feedthroughs is also required, where three are for the positioning axes. The other two feedthroughs can be used for temperature sensing, a Z-offset probe or a material feed motor.

High current feedthroughs are provided for both the material source and the substrate. The maximum expected source current, however, is only 5 A (see Section 4.4).

Positioning Stages

Three identical high resolution stepper motor stages were selected to position the substrate in a Cartesian-serial configuration. The stages are designed to be stacked, but care should be taken so that the moment acting on substrate carrying carriage is not too large. A load of 20 N can be carried by the carriage, with a maximum lateral force of 20 N and a holding force of 10 N.

The motors have two Hall-effect end stops as well as a reference position sensor. A gear box and a fine pitch lead screw achieves a 50 nm step resolution over 25 mm [201].

The stages were also custom-made on order to be vacuum compatible. This requires special lubrication and material selection. This reduces the maximum speed from 1 mm/s to 0.65 mm/s.

4.3.3 Vacuum Generation

An oil diffusion pump was available and is used to generate the the required vacuum, along with a roughing pump, valves and pressure gauges, as shown in Fig. 4.6. The use of an oil diffusion pump is further motivated by the simplicity of use, robustness and vibration-free aspects of such a pump.

The roughing pump can either evacuate the chamber through the bypass valve (V1) or the buffer through valve V2. The buffer is required to ensure that the oil diffusion pump has a sufficient pre-vacuum, when the chamber is pumped through V1 (V2 closed).

The chamber is pumped down to high vacuum through the large plate valve (V3), which connects the oil diffusion pump with the buffer. A liquid nitrogen cold trap (not shown) is also placed between V3 and the oil diffusion pump, to reduce the chance of oil droplets flowing into the chamber. A zeolite trap is placed before the roughing pump to prevent the back flow of hydrocarbons to the vacuum chamber [202].

Vibration transfer between the pumping system and the chamber is reduced by placing them on separate supports, while they are connected with a vibration damper, placed between the chamber and the large plate valve.

Four pressure gauges are used to monitor the vacuum level. Two are used to cover both the high and low vacuum ranges of the chamber, while

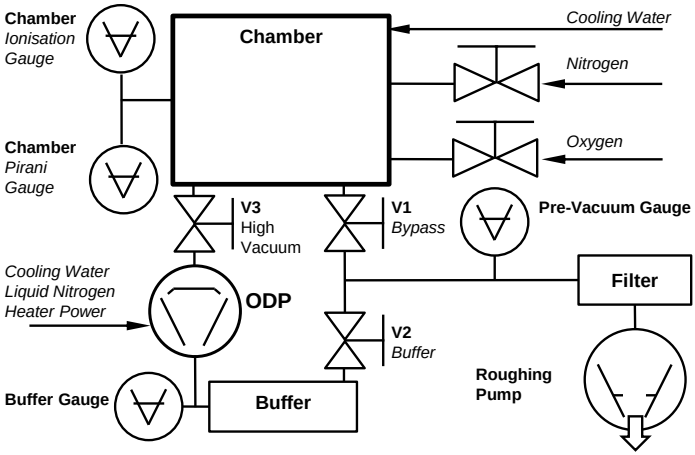


Figure 4.6: The vacuum system diagram. The oil diffusion pump is abbreviated as *ODP* and valves as *V*.

the other two are low vacuum gauges and are used to monitor the buffer and pre-vacuum pressure.

The basic system therefore only needs three valves, four gauges, cooling water and liquid nitrogen, along with the two pumps, while the pumps only require an electrical power input. The oil diffusion pump uses the electrical power to evaporate the oil, through resistive heating. The oil itself has a long service life. This is therefore an extremely robust and cost effective method to generate the required vacuum.

The automated control of the vacuum generation system was designed and implemented by Wang [202], who describes the vacuum control electronics, processes, gauges and mechanics in detail.

4.3.4 The Substrate Mount Assembly

The substrate mount assembly allows for the tilt adjustment of the substrate and it is shown in Fig. 4.7. The assembly consists of five parts, namely: (1) the arm, which is connected to the positioning stage carriage, (2) the fixed base, (3) the adjustable tilt flange, (4) the substrate mount itself and (5) a double-sided *carbon tab*.

The length of the rectangular arm component separates the heat sensitive stages from the substrate and source, while the rectangular profile reduces weight and increases stiffness. The length is determined by the shield, which is fitted around the stages and allows for the full travel of the stages to be achieved.

A cost effective substrate tilt mechanism, based on an optics mount described by Ahmad [203], is fitted to the arm. This consists of two parts,

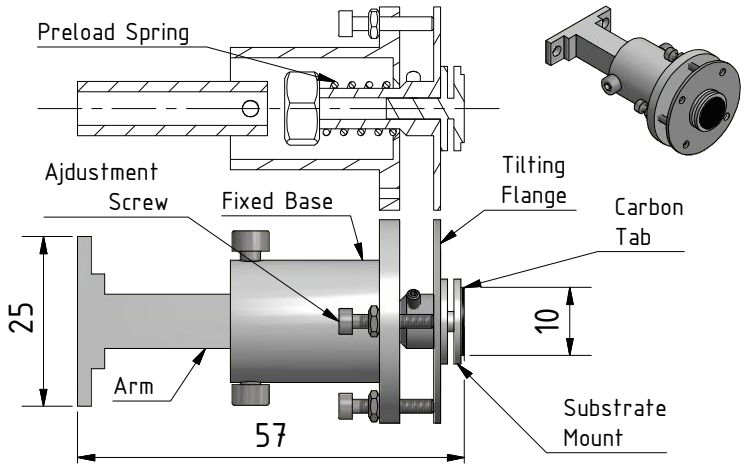


Figure 4.7: Substrate mount and tilt adjustment assembly. The cost effective tilt adjustment mechanism is based on [203].

namely the fixed base part and the flange, which can be tilted by turning the screws. In the drawing, the screws push the plate to right, while the pre-load spring forces the flange to the left.

A standard substrate mount, which is used in Scanning Electron Microscopes (SEM), is fixed to the flange. A double-sided carbon tab is used to fix the substrate to the mount. This allows for easy substrate change as well as substrate pre- or post-processing. It is also feasible to measure the deposited structure, while it is still mounted on the substrate mount.

The mass of substrate mount assembly was estimated by the CAD program as 0.026 kg. This is less than the maximum load, lateral and holding force specifications of the positioning stages (see Section 4.3.2). The system should therefore be able to move the assembly.

4.3.5 Printer Hardware and Firmware

The same firmware distribution on which the RF1000 is based, the Repetier-Firmware, is used to control the printer (see Section 2.5). Use of this firmware and the selected electronics support the *cost effective* requirement, since the work done to understand the RF1000 firmware and operation can be applied to the VDF micro-printer.

The software provides an on-line tool to configure the firmware before it can be downloaded. This can be used to set the positioning axes configuration type, main microcontroller and many other options. The VDF micro-printer, however, requires some additional modifications.

The stages have a very large range versus resolution ratio (50 nm step resolution over a 25 mm range [201]), which can make the steps per

millimetre factor large. Furthermore, the step resolution is below the default slicer positioning resolution. This also motivates the use of a faster microcontroller.

The RepRap Arduino-Due Driver Shield (RADDS) printer electronics was therefore selected. It is designed to be used with an Arduino Due which has a 32 bit ARM core. The RADDS board can, for example, drive six steppers, six heaters and three servomotors [204].

Three RADDS Power Stepper Driver (Raps128) motor driver boards, which are designed to work with RADDS shield, are used to drive the positioning stages. They can achieve up to 128 microsteps and include features such as adjustable motor decay, short circuit shut-down and adjustable motor current [204].

4.4 Source Design

The evaporation source is a critical component of the VDF concept.

This section first considers different source options and motivates why a thermal source was selected for the micro-printer. The design of the selected concept is discussed next as well as a method to estimate the source temperature. This section concludes with a simplified prediction of the heat losses from the source to the printer.

4.4.1 Source Type Selection

Different techniques which can be used to evaporate the source material are listed in Table 4.3. This includes resistive and inductive heating as well as laser ablation.

Resistive heating was selected for this project. The main reasons for this selection are the relative simplicity and the reduced safety risk associated with relatively low current resistive heating.

Here, the selection of the resistive heater material and the heat transfer mechanism to the source material are important considerations. Different heater concepts were tested to some extent, these include Surface Mount Device (SMD) resistors, high power resistors, heater wire, which is self-wound around the nozzle and the filament from a halogen light bulb.

The heater must be small, to reduce heat flow to sensitive components and the need for very high currents. This was the motivation for testing a heater using SMD resistors. Four $10\,\Omega$ resistors were placed in series on a ceramic plate (approximately $20\,\text{mm} \times 20\,\text{mm}$) and connected with silver paste. Another ceramic plate was placed on top of the resistors to increase the heat conduction away from the resistors. This was placed in a vacuum chamber at approximately 10^{-5} mbar.

The maximum power, however, which could be achieved through this method was only 10 W, before one of the resistors burnt out. The minimum heating power requirement, based on the estimated heat loss is 10 W (see also Section 4.4.4), which means that high power SMD resistors might even work in the future if the heat loss is reduced further.

Table 4.3: Heater source comparison

Type	Advantages	Disadvantages
Self-wound: Kanthal A1® wire	Low Cost	Max. temperature 1400 °C Low reproducibility Deformation of wire during heating
Self-wound: Tungsten wire	Design freedom High temperature Low Cost	Hard to wind Low reproducibility Tungsten oxidation and brittleness
Light bulb: Tungsten filament	High reproducibility High temperature Low Cost	Tungsten oxidation and brittleness Limited to what is available/design freedom
Induction Heating	Non-contact heating	More complex Physically bigger Higher currents and noise Metal source material or nozzle
Laser Ablation	Non-contact heating Any source material	Laser alignment Safety risk High cost

This shows how critical it is to know how much electrical power is needed to overcome the heat losses, which results in an iterative design cycle. An improved source can reduce the heat loss, which in-turn reduces the required heating, this can again reduce the heater size.

4.4.2 Source Mechanical Design

The heat transfer mechanisms in a vacuum environment are conduction and radiation. Heating of the source material is therefore achieved by exposing it to thermal radiation or through contact between the heat source and the material. The material, which will most likely also melt, should however not short the heater element, nor alloy with it. Indirect heating of the material is therefore desired.

The source should also reduce the required power and the heat transfer to the rest of printer.

The source design should therefore be dimensionally small and decrease any heat losses, by using thermal radiation shields and support structures with low thermal conductivity. The procurement or manufacturing of a small, high temperature, source is, however, a challenge (high temperature refers to temperatures above 1000 °C).

This is addressed here by combining existing components, which reduces the need to machine small parts from materials like stainless steel. The current resistive heating material source design is shown in Fig. 4.8.

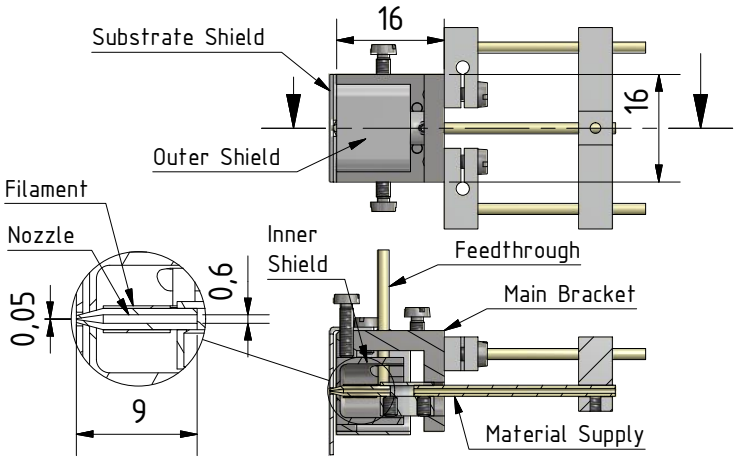


Figure 4.8: Resistive heating evaporation source.

Interestingly, a ball wire bonder capillary, which was available, has just the right dimensions and is made from alumina zirconia ceramic. This is electrically insulating and does not alloy with many PVD source materials, but does have a very low thermal conduction coefficient (depending on

grade). The exit diameter of the capillary used here is 50 µm. Smaller diameters are, however, also available.

The capillary is 9 mm long, with an external diameter of approximately 1.65 mm and an internal diameter of 0.6 mm. If a thermal expansion coefficient of $10 \times 10^{-6}\,^{\circ}\mathrm{C}^{-1}$ is assumed, then the capillary will increase in length with approximately 0.1 mm, with a 1000 °C temperature increase.

A heater wire needs to be coiled around the capillary. It is important to increase the resistance of the wire, to reduce the required current ($P = I^2R$). The resistance can be increased by reducing the cross section area and length of the wire, apart from changing the wire material itself. This is usually achieved by winding and coiling the wire.

Winding a thin, and brittle in the case of tungsten, wire around a 1.65 mm cylinder can be hard, but it is not impossible.

A standard tungsten filament light bulb (specifically, the 12 V, 50 W halogen Halostar® light bulb by OSRAM), fits tightly over the external diameter of the capillary. It is therefore used here as the heating element. This, however, does not exclude possible future experimentation with for example the Kanthal A1® wire.

Removing of the tungsten filament from the bulb is achieved by first annealing the wire (see [205]). This is done by turning it on for approximately 10 min. The glass can then be carefully broken, with for example, a bench vice and a metal box enclosure. Small pliers can be used to cut the filament legs and remove it from the fused glass. Care needs to be taken, however, as the thin tungsten filament is extremely brittle.

Electrically insulating ceramic pipes are used to feed the heating wire into the inner shield. The connection wires have hair-pin loops, which are pulled tight around the filament legs. Ideally, this is sufficient to achieve an electrical connection. Note that the filament itself is supported by the capillary. If this is not sufficient, the filament legs and the connector wire needs to be crimped or connected with silver paste, but this increases the risk of breaking the filament or evaporating the crimping material.

The evaporation material is also fed into the inner shield chamber with a ceramic pipe. A stainless steel pipe with two threaded holes is used to connect the capillary with the material feed pipe.

A high temperature ceramic paste was also investigated for either joining the heater wire to the capillary or feed pipe to the capillary. The tests were, however, not conclusive due to apparent contamination during evaporation.

4.4.3 Source Temperature Estimation

The dimensions of the enclosed source are small, which reduce the options for measuring the temperature of the source. This is further aggravated, by the need to reduce the thermal conduction losses away from the source.

The tungsten wire, however, does have a temperature dependent resistance and this, at least, can be used to estimate the temperature of the *tungsten filament*, but not of the evaporation material. It is therefore

assumed that the temperature of the filament, which is wrapped around the capillary, approximates the temperature inside the capillary, at steady state.

The relationship between the filament temperature and resistance can be determined if the room temperature resistance is known. This is required since both the length and the width of the filament are unknown, as well as the purity of the tungsten.

The room temperature resistance of the light bulb filament was measured, before removing it from the glass bulb. This is below one ohm, making it hard to measure with a hand held multi-meter.

The room temperature resistance of the filament was therefore estimated by comparing the measured resistance with the manufacturer stated operating temperature of 3000 K, at 12 V.

Three models were used and compared, namely a polynomial model cited by Stewart [206], a fitted model by Mundy [207] and a linear model $(((T - T_{ref}) * \alpha + 1) * R_0)$.

The linear model assumes a constant rate of change of resistance (α) equal to 0.0045, while R_0 was estimated as 0.216 Ω, at a reference temperature (T_{ref}) of 20 °C.

The results are shown in Fig. 4.9. Note that the two markers at approximately 2000 °C indicate the point where the power supply was changed to be able to deliver more power.

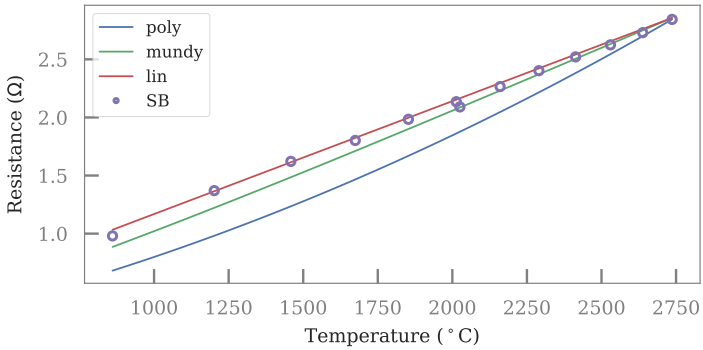


Figure 4.9: Tungsten resistance as a function of temperature, with a polynomial model (*poly*) [206], a model by *Mundy* [207], and a linear model (*lin*), compared against the temperature determined with the Stefan-Boltzmann law for the radiation (unfilled circles, *SB*) at the measured electrical input power.

The Stefan-Boltzmann law for heat radiation ($P = A\epsilon\sigma T^4$) was used to estimate the filament temperature, with Eq. (4.3), where σ is the Stefan-Boltzmann constant (taken as $5.67 \times 10^{-8}\,\mathrm{W/(m^2 \cdot K^4)}$), P the

total radiated power, which is assumed to be equal to the measured electrical power ($P = VI$) and the surface filament area approximated with a cylinder ($A = \pi DL$).

$$T = \sqrt[4]{\frac{P}{A\epsilon\sigma}} \tag{4.3}$$

The emissivity was estimated as 0.433, by using the manufacturer stated temperature of 3000 K at 50 W. The filament cylinder diameter (D) was taken as 1.6 mm (equal to the capillary external diameter), with a length of 5 mm.

The linear model appears to fit the model best. Note that effects such as the halogen gas inside the bulb and the tungsten purity were not considered. This model can, however, be used to estimate if the source can generate enough power to overcome the heat loss and evaporate the material, as discussed in Section 4.4.4, and give an indication of the source temperature.

4.4.4 Heat Loss

The heat loss of the source, especially to the positioning stages, should be minimised. This is achieved in the micro-printer by using three heat radiation shields and by reducing the thermal conduction, as shown in Fig. 4.8.

The heat flow resistance in a rod (R_{rod}) can be modelled with the electrical analogy method, given in Eq. (4.5), where L is the length of the rod, A the cross section area and k the heat conductivity of the material [208].

$$R_{rod} = \frac{L}{kA} \tag{4.4}$$

The total, steady state, heat flow ($\dot{Q}_C$) is calculated with Eq. (4.5), where ΔT is temperature difference between the rod ends. Parallel and series thermal resistances are added similar to electrical resistance networks, with the total resistance over two parallel resistances equal to $R_T = (R_a R_b)/(R_a + R_b)$ [208].

$$\dot{Q}_C = \frac{\Delta T}{R_{rod}} \tag{4.5}$$

Heat conduction losses are therefore reduced by increasing the length of the conductor, decreasing the area or by decreasing the thermal conductivity of the material. Note that a low thermal resistance can be seen as an electrical short circuit.

The heat transfer network for the material source is shown in Fig. 4.10.

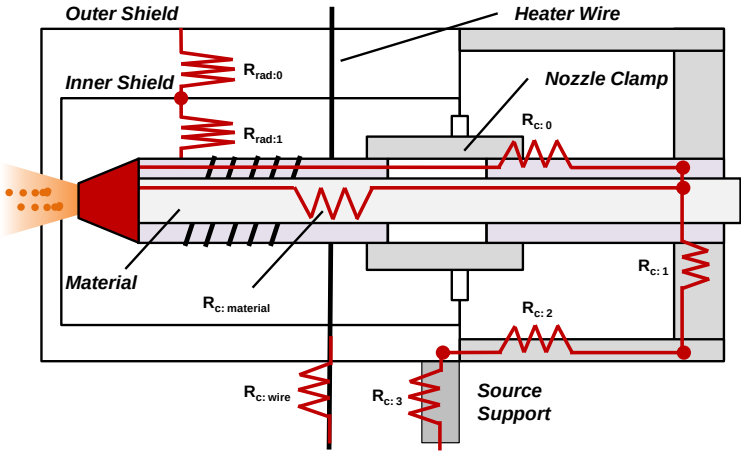


Figure 4.10: Heat loss due to conduction and radiation.

The total conduction length of the micro-printer source is increased, since the heat is first conducted through the nozzle clamp to the back of the source, then back to the source mount and finally towards the printer base plate, through the source support.

Stainless steel is used for most parts since it is vacuum compatible and has a relatively low thermal conductivity (approximately $45\,\mathrm{W \cdot K/m}$, compared to $250\,\mathrm{W \cdot K/m}$ for aluminium and $406\,\mathrm{W \cdot K/m}$ for silver).

The conductivity of the evaporation wire is also a concern. Silver is an interesting material to test the system with, since it has a relatively low evaporation temperature in high vacuum (around $1000\,^{\circ}\mathrm{C}$). This means that a 40 mm long, 0.6 mm diameter wire will have a thermal resistance of $348\,\mathrm{K/W}$ and steady state heat flow of $2.9\,\mathrm{W}$ at $1000\,^{\circ}\mathrm{C}$.

The two heater connection wires, made from tungsten, which has an approximate thermal conductivity of $173\,\mathrm{W \cdot K/m}$, will have a heat flow of $4.3\,\mathrm{W}$ at the same temperature.

The total conduction heat resistances added together results in a heat resistance for the system of $129\,\mathrm{K/W}$, which predicts a heat flow of $7.7\,\mathrm{W}$, at $1000\,^{\circ}\mathrm{C}$, for example. This means that the heater should supply more than $7.7\,\mathrm{W}$, in order to heat the source material.

Heat radiation becomes a significant form of heat transfer at higher temperatures. This is minimised by using three thermal radiation shields. Here the radiation ($\dot{Q}_r$) is approximated for the special case of two concentric cylinders and a view factor of one, with Eq. (4.6), where A is the surface area (πDL), σ the Stefan-Boltzmann constant, T the temperature

and ϵ the emissivity of each surface [208].

$$\dot{Q}_r = \frac{A_1 \sigma \left(T_1^4 - T_2^4\right)}{\frac{1}{\epsilon_1} + \frac{A_1}{A_2}\left(\frac{1}{\epsilon_2} - 1\right)} \tag{4.6}$$

The length and diameter of the inner shield are 9.5 mm and 1.65 mm, while that of the nozzle is taken as 7.5 mm and 1.65 mm, respectively. The shield is made from stainless steel, which was polished on the inside, and an emissivity of 0.16 was used for it, while 0.5 was used for the ceramic nozzle. It is also assumed that the inner shield is at room temperature, for a worse case model.

The total heat loss can now be estimated by summing the radiation and conduction heat losses, over a temperature range as shown in Fig. 4.11 (A).

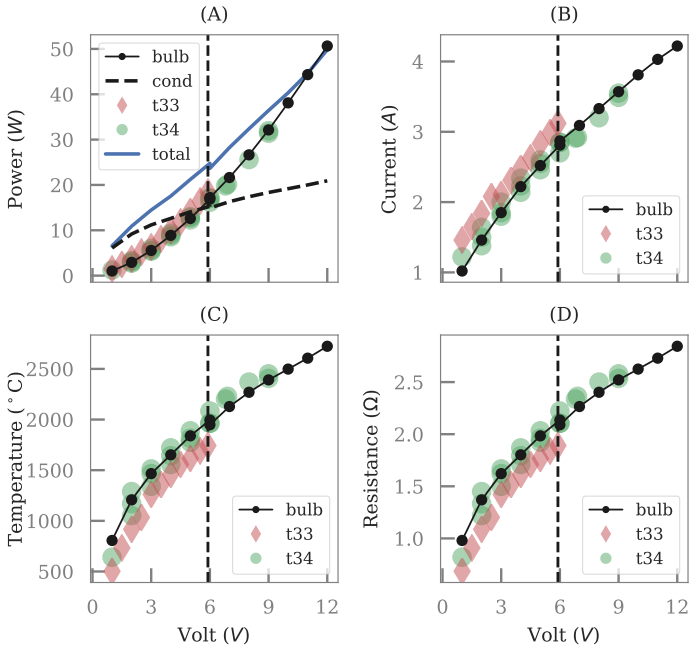


Figure 4.11: Measured and predicted heater variables against the applied voltage, for the bulb, test 33 (*t33*) and test 34 (*t34*), where (A) is the electrical power compared to the predicted total and conduction heat losses and (B-D) the current, temperature and the resistance.

The vertical dashed line indicates the point where a length of source material in test 33, which was outside the shields, starting glowing (orange-white colour). Interestingly, the heat conduction just starts to fall below the estimated heat generation at this point.

The small notch, which is also around this point, is where the voltage supply was placed in parallel operation mode to increase the power.

The increasing thermal radiation motivates the use of the shields. Test 34 used a slightly modified shield (as shown in Fig. 4.8), but did not have any source material inside the nozzle. These models motivate the minimum power requirement of 10 W.

4.5 First Tests

An initial test was performed by the micro-printer with silver wire. The test involved the following steps: (1) assembly of the printer on the main flange, (2) test the vacuum system, (3) perform a motor bake-out procedure (4) slowly heat up the source and (5) move the substrate, while the source is at a stable power level. A photo the micro-printer, after setting it up on the main flange, is shown Fig. 4.12.

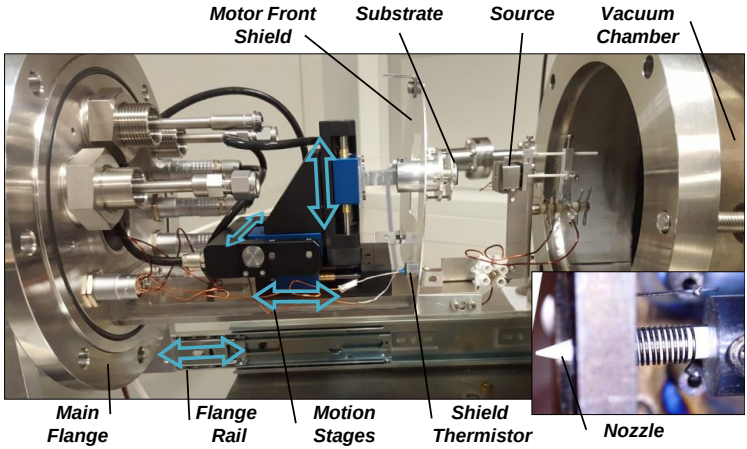


Figure 4.12: Photo of the micro-printer before the first print test, with inset showing the source assembly detail.

Two thermistors were added to monitor the temperature of the front motor shield and of one of the stages.

The vacuum pump system achieved a minimum pressure of about 5.6×10^{-6} mbar. The stages were then commanded to complete the planned positioning profile. This was repeated until the motor temperature reached approximately 35 °C.

The pressure increased to a maximum of 6.4×10^{-5} mbar during these operations. This was done to reduce the pressure increase during the actual experiment.

Photos of the source were taken during the experiment and are shown in Fig. 4.13.

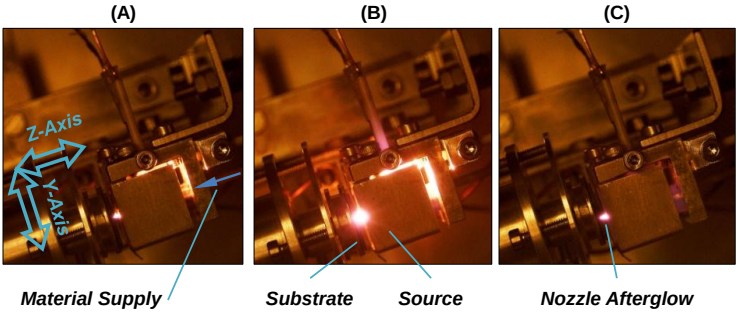


Figure 4.13: Heating the source during the first test, taken with a microscope camera through the window in the top flange. Source at approximately (A) 5 V and 2.3 A (11.5 W), (B) 7.5 V and 3.0 A (22.5 W) and (C) almost turned off.

The increasing brightness indicates a higher temperature, while the last photo shows the nozzle capillary afterglow, when the source power was almost zero. This motivates that the nozzle itself was heated up to more than 600 °C.

The first half of the G-code file was executed after increasing the source power to 8 W. This involved two lines, at different speeds, parallel to each other.The substrate is shown in Fig. 4.14.

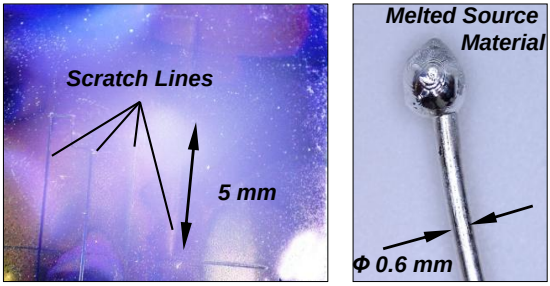


Figure 4.14: The test substrate, with the inset showing the silver source material wire, which melted inside the capillary and formed a ball at the capillary inlet.

Unfortunately, some material contaminated the substrate. Furthermore, it seems that the thermal expansion of the nozzle, could have breached the gap between the substrate and the nozzle, causing the nozzle tip to scratch the first print test substrate.

The inset in Fig. 4.14 shows that the silver material melted, and confirms that the source reached a temperature higher than 1000 °C.

5 Conclusion

The four thesis questions, which were briefly stated in the first chapter, are discussed next to conclude this thesis.

Question 1

If Additive Manufacturing (AM) is an industrial revolution, does it require new metrology?

Chapter 1 found that annual AM revenue is increasing, along with the number of patents. Furthermore, the number of journal articles in this field is also increasing. A short introduction of *Insdustry 4.0* showed how well AM fits into the paradigm of the fourth industrial revolution.

It could therefore be stated that AM is a key enabling technology for the fourth industrial revolution, rather than being an industrial revolution in itself.

Metrology is the key to change AM from a rapid prototyping tool and a complex 3D welding fabrication process into a functional component production solution. It is important to consider the four *M*'s (Market, Metrology, Materials and Make) for AM to be successful. AM also offers remarkable in-process monitoring opportunities, which gives the possibility to monitor the formation of the final object, as it is created layer by layer or particle by particle.

The trend of reducing the metrological traceability chain length was also discussed. This was considered with the modified conformity assessment diagram for AM, along with Information Rich Metrology (IRM). New metrology, which is built on the concept of Industry 4.0 cyber-physical systems, is therefore required.

This new metrology must rapidly integrate a large amount data from diverse sensor systems into process and end-product applicable information. The difference between the functional requirements and the measurement specifications must therefore also be reduced, through concurrent design for AM metrology.

Question 2

Can in-process metrology be used to improve a desktop, open source, AM Fused Filament Fabrication (FFF) process?

This work focussed on the in-process measurement and optimisation of the FFF process, specifically the filament feeding rate. A detail overview of all the main components required to achieve filament feeding and deposition were presented in Chapter 2. This included a review, with practical examples, of the motion generation tool chain and the implementation of it in the RF1000 printer. The materials and the liquefier design were also described.

Five different methods were developed in Chapter 3 to improve the FFF process, namely: (1) the optical monitoring of the feed mechanism, which can measure the volumetric flow rate, (2) a method to measure the exit flow rate, (3) a pressure sensor to measure the liquefier state, (4) single print optimisation with design of experiments and (5) a link between the in-process and post-process measured data.

The optical filament feed speed measurement method was successfully used to reduce feed slippage and active control improved the final print result. It was found that under-extrusion is a significant process factor and depends on the feed rate and the temperature.

The method was further expanded to measure the off-line exit flow rate. This was used to optimise the filament retraction settings, which was verified with a script generated test piece.

The print bed level sensor, on which the liquefier is mounted, was repurposed to measure the pressure drop over the liquefier. This was used to investigate the idea that retraction reduces the pressure on the melt and therefore reduces the ooze flow. It was, however, found that material stringing during non-print moves is also a significant factor. This shows that reducing the pressure is not the only parameter which needs to be considered, but that the dwell-time after retraction completion, nozzle tip shape and cooling are also important factors.

Single Print Optimisation (SPO) was developed so that any of the more than 100 slicer parameters can be optimised with DOE. This method drastically reduced the print time and it also presented the design of a test piece, which could be measured with affordable measurement instruments.

This paves the way towards standardisation of the testing AM machines, since the print run only requires the integrated G-code file, which is, in principle, machine independent. Researchers can therefore share this file, and do not need to modify the printer itself.

The material feed rate, filament width and the liquefier pressure were also measured during the SPO test prints. This was used to compare the in-process with the post-print measurement results with an alignment algorithm. The effect of a specific G-code on the in-process measurement results can therefore be studied with this alignment method.

The alignment algorithm was required, since the G-codes are not aware of the specific accelerations and the move joining speeds used by the

printer. This also shows the importance of the literature review, which considered the motion planning.

These methods show how an FFF printer can be monitored and optimised with open source and cost effective tools. It also paves the way for future work which can implement real time control with the pressure measurement and the standardisation of FFF measurements.

Question 3

The understanding of the FFF process was taken further to develop the Vapour Deposition Fabrication (VDF) concept. VDF is derived from dynamic stencil lithography, using PVD. The design of such a printer was presented in Chapter 4, which leads to the next question:

Is Vapour Deposition Fabrication (VDF) feasible, using Fused Filament Fabrication (FFF) as a base?

Interestingly, it was found that is possible to use FFF components (e.g. control electronics) and concepts (e.g. G-code interpreter, nozzle design, additive fabrication) to construct the VDF micro-printer. This is not only cost effective, but also develops the concept AM standardisation.

The first test showed that it is feasible, but still requires some work to reduce the substrate contamination.

Question 4

Can the answers found in Question 2 and 3 be made universally applicable to AM, by providing additional insight into Question 1?

Three answers for Question 4, based on this work, are:

- The rate of material addition in AM is a very important contributor. Measuring this rate (e.g. the exit flow rate) can be difficult. Other methods, integrated with models, are therefore required to estimate the actual material addition rate.
- Standardisation is important for AM, not only for the industrial level machines, but also for desktop printers. G-code is a good example of standardisation. The challenge is that the interpretation of the G-code should also be standardised. The benefits of standardisation are clear. An additional example is that two different processes (FFF and VDF) can potentially be controlled with the same G-code language, significantly reducing firmware development time.
- AM is a complex process with many parameters, especially due to the high resolution or the intended inhomogeneous deposition of features. Fast and reproducible optimisation is important and also possible with SPO.

List of Figures

List of Tables

References

[1] *Additive manufacturing – General principles – Terminology*. Geneva, CH: International Organization for Standardization, 2015 (pages 1, 5–7, 9, 16).

[2] T. Wohlers, I. Campbell, O. Diegel, J. Kowen, I. Fidan, and D. L. Bourell, "Wohlers Report 2017: 3D Printing and Additive Manufacturing State of the Industry Annual Worldwide Progress Report", Wohlers Associates, Inc., Tech. Rep., 2017 (pages 1–2).

[3] F. Baumann and D. Roller, "Additive Manufacturing, Cloud-Based 3D Printing and Associated Services—Overview", *Journal of Manufacturing and Materials Processing*, vol. 1, no. 2, p. 15, Oct. 2017 (pages 1–2).

[4] H. Kagermann, W. Wahlster, and J. Helbig, "Recommendations for implementing the strategic initiative INDUSTRIE 4.0", Tech. Rep. April, 2013, p. 82 (pages 2, 4).

[5] U. M. Dilberoglu, B. Gharehpapagh, U. Yaman, and M. Dolen, "The Role of Additive Manufacturing in the Era of Industry 4.0", *Procedia Manufacturing*, vol. 11, pp. 545–554, 2017 (pages 2, 4).

[6] Wikipedia. (2018). Micrometer, [Online]. Available: `https://en.wikipedia.org/w/index.php?title=Micrometer%7B%5C&%7Doldid=829723290` (visited on 03/30/2018) (page 2).

[7] International Bureau of Weights and Measures. (2018). Measurement units: the SI, [Online]. Available: `https://www.bipm.org/` (visited on 03/30/2018) (pages 2, 4).

[8] T. Doiron and J. Beers, "The Gauge Block Handbook", National Institute of Standards and Technology, Tech. Rep., 1995, p. 146 (page 2).

[9] Wenzel America. (2014). What Everybody Ought to Know About the History of CMM Coordinate Measuring Machines, [Online]. Available: `http://www.wenzelamerica.com/what-everybody-ought-to-know-about-the-history-of-cmm-coordinate-measuring-machines/` (visited on 03/30/2018) (page 4).

[10] D. Thomas and S. Gilbert, "Costs and Cost Effectiveness of Additive Manufacturing - A Literature Review and Discussion", *NIST Special Publication*, vol. 1176, pp. 1–77, 2014 (pages 4, 16).

[11] I. J. Petrick and T. W. Simpson, "3D Printing Disrupts Manufacturing", *Research Technology Management*, vol. 56, no. 6, pp. 12–16, 2013 (page 4).

[12] S. A. Tofail, E. P. Koumoulos, A. Bandyopadhyay, S. Bose, L. O'Donoghue, and C. Charitidis, "Additive manufacturing: scientific and technological challenges, market uptake and opportunities", *Materials Today*, vol. 21, no. 1, pp. 22–37, Jan. 2018 (pages 5–6, 9, 11, 15).

[13] U. Mutilba, E. Gomez-Acedo, G. Kortaberria, A. Olarra, and J. Yagüe-Fabra, "Traceability of On-Machine Tool Measurement: A Review", *Sensors*, vol. 17, no. 7, p. 1605, Jul. 2017 (pages 5, 11).

[14] S. K. Everton, M. Hirsch, P. Stravroulakis, R. K. Leach, and A. T. Clare, "Review of in-situ process monitoring and in-situ metrology for metal additive manufacturing", *Materials and Design*, vol. 95, pp. 431–445, 2016 (pages 5, 13).

[15] W. Gao, Y. Zhang, D. Ramanujan, K. Ramani, Y. Chen, C. B. Williams, C. C. Wang, Y. C. Shin, S. Zhang, and P. D. Zavattieri, "The status, challenges, and future of additive manufacturing in engineering", *Computer-Aided Design*, vol. 69, pp. 65–89, 2015 (pages 6, 15).

[16] F. Calignano, D. Manfredi, E. Ambrosio, S. Biamino, M. Lombardi, E. Atzeni, A. Salmi, P. Minetola, L. Iuliano, and P. Fino, "Overview on additive manufacturing technologies", *Proceedings of the IEEE*, vol. 105, no. 4, pp. 593–612, 2017 (pages 6, 15, 20, 55, 61).

[17] B. Redwood. (2018). Additive Manufacturing Technologies: An Overview, [Online]. Available: `https://www.3dhubs.com` (visited on 03/30/2018) (page 6).

[18] H. Kodama, "Automatic method for fabricating a three-dimensional plastic model with photo-hardening polymer", *Review of Scientific Instruments*, vol. 52, no. 11, pp. 1770–1773, 1981 (page 7).

[19] T. Wohlers and T. Gornet, "History of Additive Manufacturing 2017", Wohlers Associates, Inc., Tech. Rep., 2017, p. 40 (page 7).

[20] I. Gibson, D. W. Rosen, and B. Stucker, "Generalized Additive Manufacturing Process Chain", in *Additive Manufacturing Technologies: Rapid Prototyping to Direct Digital Manufacturing*, Boston, MA: Springer US, 2010, ch. 3, pp. 1–459 (page 7).

[21] C. Dordlofva and P. Törlind, "Qualification Challenges with Additive Manufacturing in Space Applications", *28th Annual International Solid Freeform Fabrication Symposium*, pp. 2699–2712, 2017 (pages 7, 9, 13).

[22] T. Lieneke, V. Denzer, G. A. Adam, and D. Zimmer, "Dimensional Tolerances for Additive Manufacturing: Experimental Investigation for Fused Deposition Modeling", *Procedia CIRP*, vol. 43, pp. 286–291, 2016 (page 8).

[23] I. Gibson, D. W. Rosen, and B. Stucker, "Development of Additive Manufacturing Technology", in *Additive Manufacturing Technologies: Rapid Prototyping to Direct Digital Manufacturing*, Boston, MA: Springer US, 2010, ch. 2, pp. 1–459 (page 8).

[24] T. Stock and G. Seliger, "Opportunities of Sustainable Manufacturing in Industry 4.0", *Procedia CIRP*, vol. 40, no. Icc, pp. 536–541, 2016 (page 9).

[25] J. Pellegrino, T. Makila, S. McQueen, and E. Taylor, "Measurement science roadmap for polymer-based additive manufacturing", National Institute of Standards and Technology, Gaithersburg, MD, Tech. Rep. 5, Dec. 2016 (page 10).

[26] R. Leach, P. Bointon, X. Feng, S. Lawes, S. Piano, N. Senin, D. Sims-Waterhouse, P. Stavroulakis, R. Su, W. P. Syam, and M. Thomas, "Information-rich manufacturing metrology", in *Eighth International Precision Assembly Seminar (IPAS)*, Chamonix, 2018 (pages 11, 13).

[27] "Evaluation of measurement data – The role of measurement uncertainty in conformity assessment", Joint Committee for Guides in Metrology, Tech. Rep. October, 2012, p. 57 (page 12).

[28] I. Gibson, D. W. Rosen, and B. Stucker, "Extrusion-Based Systems", in *Additive Manufacturing Technologies: Rapid Prototyping to Direct Digital Manufacturing*, Boston, MA: Springer US, 2010, ch. 6, pp. 1–459 (pages 15, 17, 22, 55).

[29] T. Wohlers and T. Gornet, "History of Additive Manufacturing 2014", *Wohlers Report 2014 - 3D Printing and Additive Manufacturing State of the Industry*, pp. 1–34, 2014 (page 15).

[30] R. Jones, P. Haufe, E. Sells, P. Iravani, V. Olliver, C. Palmer, and A. Bowyer, "Reprap - The replicating rapid prototyper", *Robotica*, vol. 29, no. 1 SPEC. ISSUE, pp. 177–191, 2011 (pages 15, 17–18, 55, 61, 68, 70).

[31] RepRap. (2018). RepRap Options, [Online]. Available: `http://reprap.org/mediawiki/index.php?title=RepRap%7B%5C_%7DOptions%7B%5C&%7Doldid=181768` (visited on 03/10/2018) (pages 15, 17, 19).

[32] A. Thompson, I. Maskery, and R. K. Leach, "X-ray computed tomography for additive manufacturing: A review", *Measurement Science and Technology*, vol. 27, no. 7, 2016 (page 15).

[33] Conrad Electronic. (2016). Renkforce RF1000 3D Drucker, [Online]. Available: `https://www.conrad.de` (visited on 09/20/2016) (pages 16, 21, 23).

[34] RepRap. (2018). RepRap G-code, [Online]. Available: `http://reprap.org/wiki/G-code` (visited on 03/12/2018) (pages 17, 34–35, 53).

[35] Wikipedia. (2018). G-code, [Online]. Available: `https://en.wikipedia.org/w/index.php?title=G-code%7B%5C&%7Doldid=829240599` (visited on 03/12/2018) (page 17).

[36] T. R. Kramer, F. M. Proctor, and E. Messina, “The NIST RS274 NGC Interpreter -Version 3”, Intelligent Systems Division, National Institute of Standards and Technology, Gaithersburg, Maryland, Tech. Rep., 2000, p. 121 (pages 17, 34–36).

[37] Arduino. (2018). Arduino, [Online]. Available: `https://www.arduino.cc` (visited on 03/15/2018) (page 18).

[38] M. Tollkühn, “Weiterentwicklung einer Ansteuerungselektronik für einen 3D-Drucker”, Bachelorarbeit, Technische Universität Braunschweig, 2014, p. 108 (pages 18, 53).

[39] J. Li, “Aufbau und Inbetriebnahme eines 3D-Druckers für den Multimaterialdruck”, Masterarbeit, Technische Universität Braunschweig, 2015 (pages 18, 20–22, 29, 69).

[40] RepRap. (2016). Glossary, [Online]. Available: `http://reprap.org/mediawiki/index.php?title=Glossary%7B%5C&%7Doldid=175471` (visited on 03/10/2018) (page 19).

[41] S. Monshausen, “Elektrisch leitfähige Filamente für den 3D-Druck”, Masterarbeit, Technische Universität Braunschweig, 2016 (pages 20, 22).

[42] G. Hodgson, A. Ranellucci, and J. Moe. (2016). Slic3r Manual - Flow Math, [Online]. Available: `http://manual.slic3r.org/advanced/flow-math` (visited on 06/21/2016) (pages 20, 90, 138).

[43] R. Jerez-Mesa, J. Travieso-Rodriguez, X. Corbella, R. Busqué, and G. Gomez-Gras, “Finite element analysis of the thermal behavior of a RepRap 3D printer liquefier”, *Mechatronics*, vol. 0, pp. 1–8, 2016 (pages 20, 88).

[44] M. Vaezi, S. Chianrabutra, B. Mellor, and S. Yang, “Multiple material additive manufacturing – Part 1: a review”, *Virtual and Physical Prototyping*, vol. 8, no. 1, pp. 19–50, 2013 (page 20).

[45] C. T. Elsworthy, “Printer head assembly for a 3D printer”, EP 2 772 347 A1, 2014 (page 22).

[46] J. Bayless, M. Chen, and B. Dai, “Wire Embedding 3D Printer”, Report, University of British Columbia, 2010 (page 22).

[47] G. T. Mark, “Methods for fiber reinforced additive manufacturing”, EP 3 130 444 A1, 2014 (page 22).

[48] C. Huber, C. Abert, F. Bruckner, M. Groenefeld, O. Muthsam, S. Schuschnigg, K. Sirak, R. Thanhoffer, I. Teliban, C. Vogler, R. Windl, and D. Suess, “3D print of polymer bonded rare-earth magnets, and 3D magnetic field scanning with an end-user 3D printer”, *Applied Physics Letters*, vol. 109, no. 16, p. 162 401, Oct. 2016 (page 22).

[49] B. Khatri, K. Lappe, D. Noetzel, K. Pursche, and T. Hanemann, “A 3D-printable polymer-metal soft-magnetic functional composite-development and characterization”, *Materials*, vol. 11, no. 2, 2018 (page 22).

[50] W. Zhou, F. A. List, C. E. Duty, and S. S. Babu, "Fabrication of conductive paths on a fused deposition modeling substrate using inkjet deposition", *Rapid Prototyping Journal*, vol. 22, no. 1, pp. 77–86, 2016 (page 22).

[51] B. Botterman, "Full Color High Definition Fused Filament Fabrication", Master's thesis, Delft Univeristy of Tecnhology, 2016 (page 22).

[52] THK. (2008). Caged Ball LM Guide SSR, [Online]. Available: `http://www.thk.com` (page 23).

[53] K. Zerfu and J. J. Ekaputri, "An approximate deflection function for simply supported quadrilateral thin plate by variational approach", in *AIP Conference Proceedings*, vol. 1867, 2017, p. 020 014 (page 24).

[54] Atmel Corporation, "AVR446: Linear speed control of stepper motor", Tech. Rep., 2006 (pages 25, 45–47).

[55] Wantai. (2018). Mini Stepper Product Specifications 42 BYGH W811 NEMA 17, [Online]. Available: `https://www.phidgets.com/productfiles/3311/3311%7B%5C_%7D0/Documentation/3311%7B%5C_%7D0%7B%5C_%7DDatasheet.pdf` (visited on 04/04/2018) (page 26).

[56] RepRap. (2016). Stepper torque, [Online]. Available: `http://reprap.org/mediawiki/index.php?title=Stepper%7B%5C_%7Dtorque%7B%5C&%7Doldid=175002` (visited on 04/11/2018) (page 26).

[57] D. W. Jones. (1998). Control of Stepping Motors: A Tutorial, [Online]. Available: `http://homepage.divms.uiowa.edu/%7B~%7Djones/step/` (visited on 04/17/2018) (pages 27, 37, 45–46).

[58] R. Condit and D. W. Jones, "Stepping Motors Fundamentals", Microchip, Tech. Rep., 2004, pp. 1–22 (pages 27, 33).

[59] Faulhaber, "Stepper Motor Technical Note: Microstepping Myths and Realities", Micromo, Faulhaber, Tech. Rep. (pages 27–28).

[60] Repetier. (2018). Repetier-Firmware Documentation, [Online]. Available: `https://www.repetier.com` (visited on 04/17/2018) (pages 28, 53).

[61] W. Bischel, "Drei-Achsen-Positionsanzeige für ein 3D-Drucksystem", Bachelorarbeit, Technische Universität Braunschweig, 2014, p. 65 (page 29).

[62] T. Hunter. (2018). Triffid Hunter's Calibration Guide, [Online]. Available: `http://reprap.org` (visited on 04/04/2018) (page 29).

[63] Thomson Industries, "Product Selection and Engineering Guide: Stepper Motor Technology", Tech. Rep. (page 30).

[64] K. S. Sollmann, M. K. Jouaneh, and D. Lavender, "Dynamic modeling of a two-axis, parallel, H-frame-type XY positioning system", *IEEE/ASME Transactions on Mechatronics*, vol. 15, no. 2, pp. 280–290, 2010 (page 30).

[65] J. Go, S. N. Schiffres, A. G. Stevens, and A. J. Hart, "Rate limits of additive manufacturing by fused filament fabrication and guidelines for high-throughput system design", *Additive Manufacturing*, vol. 16, pp. 1–11, Aug. 2017 (pages 30, 63, 67, 74, 87, 91, 94).

[66] Texas Instruments, "DRV8825 Stepper motor Controller IC", Dallas, Texas, Tech. Rep., 2014 (page 31).

[67] STMicroelectronics, "Application Note: Stepper-Motor Performance Constant-Current Chopper Drive UPS", Tech. Rep. December 2003, 2003, pp. 1–5 (page 33).

[68] F. Heinrichs, "TLE 472x – Family Stepper Motor Drivers: Current Control Method and Accuracy", Infineon Technologies AG, München,Germany, Tech. Rep. ANPS063E, 2001 (page 33).

[69] Reprap. (2018). List of Firmware, [Online]. Available: `http://reprap.org` (visited on 04/07/2018) (page 35).

[70] GitHub. (2017). About Grbl, [Online]. Available: `https://github.com/gnea/grbl/wiki` (visited on 04/07/2018) (pages 35, 51).

[71] D. Cary. (2015). Firmware: Linear Acceleration, [Online]. Available: `http://reprap.org` (visited on 04/05/2018) (pages 38–39).

[72] Repetier. (Oct. 2011). Hardware settings and print quality, [Online]. Available: `https://github.com/repetier/Repetier-Firmware/wiki/Hardware-settings-and-print-quality` (visited on 09/19/2016) (pages 39–40, 53, 83–84, 145).

[73] J. Sungeun. (2011). Improving Grbl: Cornering Algorithm, [Online]. Available: `https://onehossshay.wordpress.com/2011/09/24/improving%7B%5C_%7Dgrbl%7B%5C_%7Dcornering%7B%5C_%7Dalgorithm/` (visited on 04/17/2018) (page 40).

[74] H. M. Traumflug. (2016). Printer Controller Trajectory Planning, [Online]. Available: `https://www.reprap-diy.com/printer%7B%5C_%7Dcontroller%7B%5C_%7Dtrajectory%7B%5C_%7Dplanning` (visited on 04/05/2018) (page 42).

[75] J. I. Quinones, "Applying acceleration and deceleration profiles to bipolar stepper motors", Texas Instruments Incorporated, Tech. Rep., 2012, p. 7 (page 42).

[76] J. Q. Oberhauser, "Design, Construction, Control, and Analysis of Linear Delta Robot", Master's thesis, Ohio University, 2016 (page 43).

[77] R. Giseburt. (2016). Jerk Controlled Motion Explained, [Online]. Available: `https://github.com/synthetos/TinyG/wiki/Jerk-Controlled-Motion-Explained` (visited on 04/05/2018) (page 43).

[78] D. Austin, "Generate stepper-motor speed profiles in real time", Embedded, Tech. Rep., 2004, pp. 1–5 (pages 45–47).

[79] B. Weiss, D. W. Storti, and M. A. Ganter, "Low-cost closed-loop control of a 3D printer gantry", *Rapid Prototyping Journal*, vol. 21, no. 5, D. Eujin Pei, Ed., pp. 482–490, Aug. 2015 (pages 45, 91, 93).

[80] P. Ranade, "Linear motor control without the math", *EE Times*, 2009 (pages 45, 48).

[81] PicProg. (2012). Stepping Motion Profiles in Realtime, [Online]. Available: `http://picprog.strongedge.net/step%7B%5C_%7Dprof/step-profile.html` (visited on 04/19/2018) (page 45).

[82] J. E. Bresenham, "Algorithm for computer control of a digital plotter", *IBM Systems Journal*, vol. 4, no. 1, pp. 25–30, 1965 (page 51).

[83] K. I. Joy, "Breshenham's Algorithm", Visualization and Graphics Research Group,Department of Computer Science, Davis, Tech. Rep., 1999, pp. 1–15 (page 51).

[84] B. Hampel, S. Monshausen, and M. Schilling, "Properties and applications of electrically conductive thermoplastics for additive manufacturing of sensors", *Technisches Messen*, vol. 84, no. 9, pp. 593–599, 2017 (page 55).

[85] H. Giberti, L. Sbaglia, and M. Silvestri, "Mechatronic Design for an Extrusion-Based Additive Manufacturing Machine", *Machines*, vol. 5, no. 4, p. 29, Nov. 2017 (pages 55, 61, 68).

[86] S. Rohringer. (2018). 3D Printer Filament Guide 2018 – The Top 25 Types, [Online]. Available: `https://all3dp.com/1/3d-printer-filament-types-3d-printing-3d-filament/` (visited on 04/18/2018) (page 55).

[87] RigidInk. (2018). Complete 3D Printing Filament Comparison Guide, [Online]. Available: `https://rigid.ink/pages/filament-comparison-guide` (visited on 04/18/2018) (page 55).

[88] Simplify3d. (2018). Filament Properties Table, [Online]. Available: `https://www.simplify3d.com/support/materials-guide/properties-table/` (visited on 04/18/2018) (pages 55, 58–60).

[89] M. Diagne and M. Krstic, "State dependent input delay compensated Bang Bang control: Application to 3D printing based on screw extruder", *Proceedings of the American Control Conference*, vol. 2015-July, pp. 5653–5658, 2015 (page 61).

[90] B. Stephens, P. Azimi, Z. El Orch, and T. Ramos, "Ultrafine particle emissions from desktop 3D printers", *Atmospheric Environment*, vol. 79, pp. 334–339, 2013 (page 61).

[91] P. Azimi, D. Zhao, C. Pouzet, N. E. Crain, and B. Stephens, "Emissions of Ultrafine Particles and Volatile Organic Compounds from Commercially Available Desktop Three-Dimensional Printers with Multiple Filaments", *Environmental Science and Technology*, vol. 50, no. 3, pp. 1260–1268, 2016 (page 61).

[92] L. Mendes, A. Kangas, K. Kukko, B. Mølgaard, A. Säämänen, T. Kanerva, I. Flores Ituarte, M. Huhtiniemi, H. Stockmann-Juvala, J. Partanen, K. Hämeri, K. Eleftheriadis, and A. K. Viitanen, "Characterization of Emissions from a Desktop 3D Printer", *Journal of Industrial Ecology*, vol. 21, S94–S106, 2017 (page 61).

[93] L. T. Sin, A. R. Rahmat, and W. A. W. A. Rahman, "Thermal Properties of Poly(lactic Acid)", in *Polylactic Acid*, 5, vol. 52, Elsevier, 2013, pp. 109–141 (page 62).

[94] D. Henton, P. Gruber, J. Lunt, and J. Randall, "Polylactic Acid Technology", in *Natural Fibers, Biopolymers, and Biocomposites*, CRC Press, Apr. 2005, pp. 527–578 (page 62).

[95] NatureWorks, *Ingeo Biopolymer 3D850 Technical Data Sheet: 3D Printing Monofilament – High Heat Grade*, Minnetonka, USA, 2018 (pages 62–63).

[96] J. W. Comb, W. R. Priedeman, and P. W. Turley, "FDM technology process improvements", *Solid Freeform Fabrication Proceedings*, pp. 42–49, 1994 (pages 62, 77, 94).

[97] J. Vlachopoulos and D. Strutt, "The role of rheology in polymer extrusion", in *New Technology for Extrusion*, Milan, Italy, Nov. 2003 (pages 62–64).

[98] Q. Fang and M. A. Hanna, "Rheological properties of amorphous and semicrystalline polylactic acid polymers", *Industrial Crops and Products*, vol. 10, no. 1, pp. 47–53, Jun. 1999 (pages 62–64, 66–67).

[99] J. R. Dorgan, "Rheology of poly(lactic acid)", in *Poly(lactic acid): Synthesis, Structures, Properties, Processing, and Applications*, New Jersey: John Wiley & Sons, Inc., 2010, ch. 10, pp. 125–139 (pages 62–64).

[100] J. Lai, "Moldflow Material Testing Report: MAT2238 NatureWorks PLA", Moldflow Plastics Labs, Victoria,Australia, Tech. Rep., 2007 (pages 63–66, 74, 79, 84, 173).

[101] S. Dabiri, S. Schmid, and G. Tryggvason, "Fully Resolved Numerical Simulations of Fused Deposition Modeling", *Volume 2: Processing*, vol. 2, V002T02A045, 2014 (pages 63, 88).

[102] A. Bellini, S. Güçeri, and M. Bertoldi, "Liquefier Dynamics in Fused Deposition", *Journal of Manufacturing Science and Engineering*, vol. 126, no. 2, p. 237, 2004 (pages 63, 79–84, 86–88, 91, 173).

[103] H. S. Ramanath, C. K. Chua, K. F. Leong, and K. D. Shah, "Melt flow behaviour of PCL in fused deposition modelling", *Journal of Materials Science: Materials in Medicine*, vol. 19, no. 7, pp. 2541–2550, 2008 (pages 63, 80, 88).

[104] N. Mostafa, H. M. Syed, S. Igor, and G. Andrew, "A study of melt flow analysis of an ABS-Iron composite in fused deposition modelling process", *Tsinghua Science and Technology*, vol. 14, no. S1, pp. 29–37, Jun. 2009 (pages 63, 88).

[105] M. Roxas, S. Ju, and P. E. Sullivan, "Fluid Dynamics Analysis of Desktop-based Fused Deposition Modeling Rapid Prototyping", Bachelor, University of Toronto, 2008 (pages 63, 80, 173).

[106] T. Hofstätter, D. B. Pedersen, J. S. Nielsen, R. Pimentel, M. Mischkot, and H. N. Hansen, "Calibration of a numerical model for heat transfer and fluid flow in an extruder", *International Journal of Rapid Manufacturing*, vol. 6, no. 1, pp. 1–16, 2016 (pages 63, 91).

[107] K. Hamad, M. Kaseem, H. W. Yang, F. Deri, and Y. G. Ko, "Properties and medical applications of polylactic acid: A review", *Express Polymer Letters*, vol. 9, no. 5, pp. 435–455, 2015 (pages 63–64).

[108] L. T. Sin, A. R. Rahmat, and W. A. W. A. Rahman, "Rheological Properties of Poly(lactic Acid)", in *Polylactic Acid*, Elsevier, 2013, pp. 221–245 (pages 64–65).

[109] S. Djellali, T. Sadoun, N. Haddaoui, and A. Bergeret, "Viscosity and viscoelasticity measurements of low density polyethylene/poly (lactic acid) blends", *Polymer Bulletin*, vol. 72, no. 5, pp. 1177–1195, May 2015 (page 64).

[110] T. A. Osswald and N. Rudolph, *Polymer Rheology*. Munich: Hanser Publishers, 2015, p. 225 (pages 64–66).

[111] N. Venkataraman, S. Rangarajan, M. Matthewson, B. Harper, A. Safari, S. Danforth, G. Wu, N. Langrana, S. Guceri, and A. Yardimci, "Feedstock material property – process relationships in fused deposition of ceramics (FDC)", *Rapid Prototyping Journal*, vol. 6, no. 4, pp. 244–253, 2000 (page 65).

[112] Q.-H. Nguyen and N.-D. Nguyen, "Incompressible Non-Newtonian Fluid Flows", in *Continuum Mechanics - Progress in Fundamentals and Engineering Applications*, InTech, Mar. 2012, pp. 47–72 (pages 66, 79).

[113] H. Li, T. Wang, J. Sun, and Z. Yu, "The effect of process parameters in fused deposition modelling on bonding degree and mechanical properties", *Rapid Prototyping Journal*, vol. 24, no. 1, pp. 80–92, Jan. 2018 (pages 68, 94–95).

[114] R. S. Crockett and P. D. Calvert, "The Liquid-to-Solid Transition in Stereodeposition Techniques", in *Solid Freeform Fabrication Proceedings*, University of Texas, 1996, pp. 257–264 (page 68).

[115] B. N. Turner, R. Strong, and S. A. Gold, "A review of melt extrusion additive manufacturing processes: I. Process design and modeling", *Rapid Prototyping Journal*, vol. 20, no. 3, pp. 192–204, Apr. 2014 (pages 68, 72–74, 77, 80, 84–85, 115).

[116] D. Ahn, J. H. Kweon, S. Kwon, J. Song, and S. Lee, "Representation of surface roughness in fused deposition modeling", *Journal of Materials Processing Technology*, vol. 209, no. 15-16, pp. 5593–5600, 2009 (page 68).

[117] A. Boschetto and L. Bottini, "Accuracy prediction in fused deposition modeling", *International Journal of Advanced Manufacturing Technology*, vol. 73, no. 5-8, pp. 913–928, 2014 (pages 68, 95, 145).

[118] M. Temming, “Aufbau und Inbetriebnahme eines Pasten- extruders für einen 3D-Drucker”, Bachelorarbeit, Technische Universität Braunschweig, 2014 (page 68).

[119] C. Amza, A. Zapciu, and D. Popescu, “Paste Extruder—Hardware Add-On for Desktop 3D Printers”, *Technologies*, vol. 5, no. 3, p. 50, 2017 (page 68).

[120] D. Drotman, M. Diagne, R. Bitmead, and M. Krstic, “Control-oriented energy-based modeling of a screw extruder used for 3D printing”, in *ASME 2016 Dynamic Systems and Control Conference, DSCC 2016*, vol. 2, Minneapolis, Minnesota, USA, 2016, p. 31 (pages 68–69).

[121] H. Valkenaers, F. Vogeler, E. Ferraris, A. Voet, and J.-P. Kruth, “A novel approach to additive manufacturing: screw extrusion 3D-printing”, in *4M Conference*, 2013, pp. 235–238 (page 68).

[122] R. L. Zinniel and J. S. Batchelder, “Conical viscosity pump with axially positionable impeller and method of printing a 3D part.”, US 2017 012 23 22 A1, 2017 (page 68).

[123] J. S. Batchelder and W. J. Swanson, “Filament drive mechanism for use in extrusion-based digital manufacturing systems”, US 7 896 209 B2, 2011 (page 70).

[124] J. S. Batchelder and R. R. Jackson, “Method and Apparatus for Solid Prototyping”, US 5 764 521, 1998 (page 70).

[125] M. Fiedler, “Evaluating Tension and Tooth Geometry to Optimize Grip on 3D Printer Filament”, *3D Printing and Additive Manufacturing*, vol. 2, no. 2, pp. 85–88, Jun. 2015 (page 72).

[126] W. J. Heij, “Feed velocity feedback for high speed fused deposition modelling machines”, M. Sc. Delft Univeristy of Tecnhology, 2016 (pages 73, 83, 86–88, 91, 98).

[127] M. a. Yardimci, T. Hattori, S. I. Guceri, and S. C. Danforth, “Thermal analysis of Fused Deposition”, *Solid Freeform Fabrication Proceedings, September 1997*, pp. 689–698, 1997 (pages 74, 85).

[128] Torwell Australia. (2018). Duraplat Coated Nozzles, [Online]. Available: `https://www.torwell3d.com.au/` (visited on 04/27/2018) (pages 74–76).

[129] E3D Blog. (2016). Silicone Socks: Cleaner Nozzles And Sharper Prints, [Online]. Available: `https://e3d-online.com/blog/2016/08/13/silicone-socks/` (visited on 04/27/2018) (page 76).

[130] J. Francois. (2016). 3D printing nozzle characteristics, [Online]. Available: `http://www.tridimake.com/2016/06/3d-printing-nozzles-characteristics.html` (visited on 04/27/2018) (pages 76–77).

[131] W. J. Swanson, D. F. Mannella, K. C. Johnson, and R. G. Schloesser, “Print Head Nozzle for Use With Additive Manufacturing System”, US 2014 004 89 69 A1, 2014 (page 76).

[132] Jerrill. (2013). What's in a nozzle?, [Online]. Available: http://www.jerrill.com/blog/2013/04/whats-in-a-nozzle/ (visited on 04/27/2018) (page 77).

[133] C. McIlroy and P. D. Olmsted, "Deformation of an amorphous polymer during the fused-filament-fabrication method for additive manufacturing", *Journal of Rheology*, vol. 61, no. 2, pp. 379–397, Mar. 2017 (page 77).

[134] N. A. Sukindar and M. K. A. Mohd Ariffin, "An Analysis on Finding the Optimum Die Angle of Polylactic Acid in Fused Deposition Modelling", *Applied Mechanics and Materials*, vol. 835, pp. 254–259, 2016 (pages 77, 88).

[135] Wikipedia. (2018). Hagen–Poiseuille equation, [Online]. Available: https://en.wikipedia.org/w/index.php?title=Hagen%E2%80%93Poiseuille%7B%5C_%7Dequation%7B%5C&%7Doldid=832771843 (visited on 04/28/2018) (page 77).

[136] T. Sochi, "Newtonian Flow in Converging-Diverging Capillaries", *International Journal of Modeling, Simulation, and Scientific Computing*, vol. 04, no. 03, p. 1 350 011, Sep. 2013 (pages 78–79).

[137] W. Michaeli, *Extrusion Dies for Plastics and Rubber: Design and Engineering Computation*, Second Edi. New York: Hanser ed., 1992 (page 79).

[138] H. Bikas, P. Stavropoulos, and G. Chryssolouris, "Additive manufacturing methods and modelling approaches: a critical review", *The International Journal of Advanced Manufacturing Technology*, vol. 83, no. 1-4, pp. 389–405, Mar. 2016 (pages 80, 173).

[139] Comb, "Liquifier pump control in an extrusion apparatus", US 6814907, 2004 (pages 82–83, 118).

[140] M. Roberts. (2011). Mattroberts' Firmware, [Online]. Available: http://reprap.org/mediawiki/index.php?title=Mattroberts%7B%5C%%7D27%7B%5C_%7DFirmware%7B%5C&%7Doldid=47350 (visited on 04/30/2018) (page 83).

[141] M. Henschke. (2017). Advance-Algorithmus, [Online]. Available: http://www.dr-henschke.de/waerme%7B%5C_%7Dfilament.html (visited on 04/30/2018) (page 84).

[142] J. S. Batchelder, W. J. Swanson, and K. C. Johnson, "Additive Manufacturing System and Process with Material Flow Feedback Control", US 2015 009 73 08 A1, 2013 (pages 85, 88, 92).

[143] J. Go and A. J. Hart, "Fast Desktop-Scale Extrusion Additive Manufacturing", *Additive Manufacturing*, vol. 18, pp. 276–284, Dec. 2017 (pages 86, 94).

[144] M. K. Agarwala, V. R. Jamalabad, N. a. Langrana, A. Safari, P. J. Whalen, and S. C. Danforth, "Structural quality of parts processed by fused deposition", *Rapid Prototyping Journal*, vol. 2, no. 4, pp. 4–19, Dec. 1996 (page 86).

[145] G. P. Greeff and M. Schilling, "Closed loop control of slippage during filament transport in molten material extrusion", *Additive Manufacturing*, vol. 14, pp. 31–38, 2017 (pages 86, 91, 93, 98, 116, 119, 121–122).

[146] R. Comminal, M. P. Serdeczny, D. B. Pedersen, and J. Spangenberg, "Numerical modeling of the strand deposition flow in extrusion-based additive manufacturing", *Additive Manufacturing*, vol. 20, no. March, pp. 68–76, 2018 (pages 88–91, 168).

[147] R. L. Zinniel and J. S. Batchelder, "Volumetric Feed Control for Flexible Filament", US 6 085 957, 2000 (page 92).

[148] J. S. Batchelder, "Additive Manufacturing System and Method for Printing Three-Dimensional Parts Using Velocimetry", US 13 840 538, 2014 (page 92).

[149] D. J. Hoelzle, A. G. Alleyne, and A. J. Wagoner Johnson, "Iterative Learning Control for robotic deposition using machine vision", in *2008 American Control Conference*, IEEE, Jun. 2008, pp. 4541–4547 (page 92).

[150] M. Faes, W. Abbeloos, F. Vogeler, C. K, H. Valkenaers, T. Goedeme, and E. Ferraris, "Process Monitoring of Extrusion Based 3D Printing via Laser Scanning", in *International Conference on Polymers and Moulds Innovations - PMI2014*, vol. 6, 2014 (page 92).

[151] D. Pollard, C. Ward, G. Herrmann, and J. Etches, "Filament Temperature Dynamics in Fused Deposition Modelling and Outlook for Control", *Procedia Manufacturing*, vol. 11, no. June, pp. 536–544, 2017 (page 92).

[152] I. Alig, B. Steinhoff, and D. Lellinger, "Monitoring of polymer melt processing", *Measurement Science and Technology*, vol. 21, no. 6, pp. 1–19, 2010 (page 92).

[153] T. Paul and J. Batchelder, "Capacitive detector for use in extrusion-based digital manufacturing systems", US 8 222 908 B2, 2012 (page 92).

[154] RADDS. (2018). FTS – Filament Tracking System, [Online]. Available: `http://doku.radds.org` (visited on 03/21/2018) (page 92).

[155] C. Kim, D. Espalin, A. Cuaron, M. A. Perez, E. MacDonald, and R. B. Wicker, "A study to detect a material deposition status in fused deposition modeling technology", in *2015 IEEE International Conference on Advanced Intelligent Mechatronics (AIM)*, vol. 2015-Augus, IEEE, Jul. 2015, pp. 779–783 (page 92).

[156] Y. Tlegenov, Y. S. Wong, and G. S. Hong, "A dynamic model for nozzle clog monitoring in fused deposition modelling", *Rapid Prototyping Journal*, vol. 23, no. 2, pp. 391–400, Mar. 2017 (page 93).

[157] S. Bukkapatnam and B. Clark, "Dynamic Modeling and Monitoring of Contour Crafting—An Extrusion-Based Layered Manufacturing Process", *Journal of Manufacturing Science and Engineering*, vol. 129, no. 1, p. 135, 2007 (page 93).

[158] H. Wu, Y. Wang, and Z. Yu, "In situ monitoring of FDM machine condition via acoustic emission", *The International Journal of Advanced Manufacturing Technology*, Sep. 2015 (page 93).

[159] P. K. Rao, J. P. Liu, D. Roberson, Z. J. Kong, and C. Williams, "Online Real-Time Quality Monitoring in Additive Manufacturing Processes Using Heterogeneous Sensors", *Journal of Manufacturing Science and Engineering*, vol. 137, no. 6, p. 061 007, Sep. 2015 (page 93).

[160] F. Baumann, M. Schön, J. Eichhoff, and D. Roller, "Concept Development of a Sensor Array for 3D Printer", *Procedia CIRP*, vol. 51, pp. 24–31, 2016 (page 93).

[161] T. Fang, M. A. Jafari, S. C. Danforth, and A. Safari, "Signature analysis and defect detection in layered manufacturing of ceramic sensors and actuators", *Machine Vision and Applications*, vol. 15, no. 2, pp. 63–75, 2003 (page 93).

[162] F. Baumann and D. Roller, "Vision based error detection for 3D printing processes", vol. 03, pp. 3–9, 2016 (page 93).

[163] J. Straub, "Initial Work on the Characterization of Additive Manufacturing (3D Printing) Using Software Image Analysis", *Machines*, vol. 3, no. 2, pp. 55–71, Apr. 2015 (page 93).

[164] S. Nuchitprasitchai, M. Roggemann, and J. Pearce, "Three Hundred and Sixty Degree Real-Time Monitoring of 3-D Printing Using Computer Analysis of Two Camera Views", *Journal of Manufacturing and Materials Processing*, vol. 1, no. 1, p. 2, Jul. 2017 (page 93).

[165] H. Sun, P. K. Rao, Z. Kong, X. Deng, and R. Jin, "Functional Quantitative and Qualitative Models for Quality Modeling in a Fused Deposition Modeling Process", *IEEE Transactions on Automation Science and Engineering*, pp. 1–11, 2017 (pages 94–95, 145).

[166] A. Peng, X. Xiao, and R. Yue, "Process parameter optimization for fused deposition modeling using response surface methodology combined with fuzzy inference system", *The International Journal of Advanced Manufacturing Technology*, vol. 73, no. 1-4, pp. 87–100, 2014 (page 94).

[167] O. A. Mohamed, S. H. Masood, and J. L. Bhowmik, "Optimization of fused deposition modeling process parameters: a review of current research and future prospects", *Advances in Manufacturing*, vol. 3, no. 1, pp. 42–53, Mar. 2015 (page 94).

[168] H. Xia and J. Lu, "Fully Resolved Numerical Simulations of Fused Deposition Modeling . Part II — Solidification , Residual Stresses , and Modeling of the Nozzle", pp. 1–24, 2017 (page 94).

[169] M. Shahrain, T. Didier, G. K. Lim, and A. J. Qureshi, "Fast Deviation Simulation for 'Fused Deposition Modeling' Process", *Procedia CIRP*, vol. 43, pp. 327–332, 2016 (pages 95, 145).

[170] Repetier. (2016). Repetier-Server API description, [Online]. Available: https://www.repetier-server.com/manuals/programming/API/index.html (visited on 09/20/2016) (page 100).

[171] Crossbar.io. (2016). Open-source real-time framework for Web, Mobile & Internet of Things, [Online]. Available: http://autobahn.ws/ (visited on 09/20/2016) (page 100).

[172] G. Bradski, "The OpenCV Library", *Dr. Dobb's Journal of Software Tools*, 2000 (page 100).

[173] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision.", in *IJCAI*, vol. 81, 1981, pp. 674–679 (pages 101, 103).

[174] J.-Y. Bouguet, "Pyramidal implementation of the affine Lucas Kanade feature tracker description of the algorithm", *Intel Corporation*, vol. 5, no. 1-10, p. 4, 2001 (page 101).

[175] J. E. Solem, *Programming Computer Vision with Python*, Final Draft. 2012, p. 264 (pages 102–103).

[176] Wikipedia. (2018). Optical Flow, [Online]. Available: https://en.wikipedia.org/w/index.php?title=Optical%7B%5C_%7Dflow%7B%5C&%7Doldid=842522656 (visited on 05/26/2018) (page 103).

[177] A. Mordvintsev and K. Abid. (2013). Optical Flow Tutorial, [Online]. Available: http://opencv-python-tutroals.readthedocs.io/en/latest/py%7B%5C_%7Dtutorials/py%7B%5C_%7Dvideo/py%7B%5C_%7Dlucas%7B%5C_%7Dkanade/py%7B%5C_%7Dlucas%7B%5C_%7Dkanade.html (visited on 10/26/2016) (pages 103, 105).

[178] E. Jones, T. Oliphant, P. Peterson, *et al.* (2000). SciPy: Open Source Scientific Tools for Python, [Online]. Available: http://www.scipy.org/ (visited on 05/08/2018) (pages 108, 117, 173, 178).

[179] "GUM Workbench: The tool for the Expression of Uncertainty in Measurement", Metrodata GmbH, Weil am Rhein, Germany, Tech. Rep., 2008 (page 110).

[180] G. P. Greeff and M. Schilling, "Comparing Retraction Methods with Volumetric Exit Flow Measurement in Molten Material Extrusion", in *Special Interest Group meeting on Dimensional Accuracy and Surface Finish in Additive Manufacturing*, W. Dewulf, H. N. Hansen, R. K. Leach, J. Qian, and J. S. Taylor, Eds., Leuven, Belgium: euspen, 2017, pp. 70–74 (page 123).

[181] N. C. Lewis. (2017). The Arduino InkShield: a HP C6602 inkjet cartridge controller, [Online]. Available: http://nicholasclewis.com/projects/inkshield/ (visited on 06/30/2017) (page 124).

[182] G. P. Greeff and M. Schilling, "Single print optimisation of fused filament fabrication parameters", *The International Journal of Advanced Manufacturing Technology*, Aug. 2018 (page 138).

[183] NIST/SEMATECH. (2013). e-Handbook of Statistical Methods, [Online]. Available: http://www.itl.nist.gov/div898/handbook/ (visited on 03/21/2018) (page 141).

[184] S. Seabold and J. Perktold, "Statsmodels: Econometric and statistical modeling with python.", in *Proceedings of the 9th Python in Science Conference*, 2010 (page 141).

[185] G. A. O. Adam and D. Zimmer, "On design for additive manufacturing: Evaluating geometrical limitations", *Rapid Prototyping Journal*, vol. 21, no. 6, pp. 662–670, 2015 (page 143).

[186] S. Moylan, J. Slotwinski, A. Cooke, K. Jurrens, M. A. Donmez, and A. Donmez, "Proposal for a Standardized Test Artifact for Additive Manufacturing Machines and Processes", *Solid Freeform Fabrication Symposium Proceedings*, pp. 902–920, 2012 (page 145).

[187] D. Norée. (2018). Benchy: the jolly 3D printing torture-test, [Online]. Available: http://www.3dbenchy.com/ (visited on 03/21/2018) (page 145).

[188] V. Savu, M. A. F. van den Boogaart, J. Brugger, J. Arcamone, M. Sansa, and F. Perez-Murano, "Dynamic stencil lithography on full wafer scale", *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures*, vol. 26, no. 6, pp. 2054–2058, Nov. 2008 (pages 186, 188–189).

[189] N. Jacobson, "The Diffusion of Gases in Capillaries", Ph.D. thesis, University of California, 1981 (page 186).

[190] D. M. Mattox, "Vacuum Evaporation and Vacuum Deposition", in *Handbook of Physical Vapor Deposition (PVD) Processing*, Elsevier, 2010, pp. 195–235 (page 187).

[191] T. E. Dunham and J. P. Hirth, "Effusion from Knudsen Cells with Conical Channels", *The Journal of Chemical Physics*, vol. 49, no. 10, pp. 4650–4659, Nov. 1968 (page 187).

[192] R. Hołyst, M. Litniewski, and D. Jakubczyk, "A molecular dynamics test of the Hertz–Knudsen equation for evaporating liquids", *Soft Matter*, vol. 11, no. 36, pp. 7201–7206, 2015 (page 187).

[193] K. Robbie and M. J. Brett, "Sculptured thin films and glancing angle deposition: Growth mechanics and applications", *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films*, vol. 15, no. 3, pp. 1460–1465, May 1997 (page 188).

[194] Y. Zhao, "Dynamic Shadowing Growth and Its Energy Applications", *Frontiers in Energy Research*, vol. 2, no. September, pp. 1–8, Sep. 2014 (page 188).

[195] O. Vazquez-Mena, L. Gross, S. Xie, L. Villanueva, and J. Brugger, "Resistless nanofabrication by stencil lithography: A review", *Microelectronic Engineering*, vol. 132, pp. 236–254, Jan. 2015 (pages 188–189).

[196] C. Cojocaru, C. Harnagea, A. Pignolet, and F. Rosei, "Patterning of Functional Materials by Pulsed Laser Deposition through Nanostencils", in *2006 IEEE Conference on Emerging Technologies - Nanoelectronics*, vol. 2006, IEEE, 2006, pp. 283–288 (page 188).

[197] T. Schallenberg, L. W. Molenkamp, and G. Karczewski, "Molecular beam epitaxy of compound materials through shadow masks", *Physical Review B*, vol. 70, no. 15, p. 155 328, Oct. 2004 (page 188).

[198] V. Savu, S. Xie, and J. Brugger, "100 mm dynamic stencils pattern sub-micrometre structures", *Nanoscale*, vol. 3, no. 7, p. 2739, 2011 (pages 188–189).

[199] F. T. Wallenberger, "Rapid Prototyping Directly from the Vapor Phase", *Science*, vol. 267, no. 5202, pp. 1274–1275, 1995 (page 190).

[200] D. Coyne, "LIGO Vacuum Compatible Materials", Laser Interferometer Gravitational Wave Observatory, Tech. Rep., 2011 (page 193).

[201] P. I. (GmbH. (2018). Compact Micro-Translation Stage, [Online]. Available: `https://www.physikinstrumente.com/en/products/linear-stages-and-actuators/stages-with-motor-screw-drives/m-110-m-111-m-112-compact-micro-translation-stage-701650/` (visited on 06/04/2018) (pages 196, 198).

[202] S. Wang, "Automatic Control of a Vacuum System", Masterarbeit, Technische Universität Braunschweig, 2018 (pages 196–197).

[203] A. Ahmad, "Adjustment Mechanisms", in *Optomechanical Engineering Handbook*, A. Ahmad, Ed., Florida, USA: CRC Press, 1999, ch. Chapter 7 (pages 197–198).

[204] Wordpress. (2016). RADDS Electronics for 3D Printer – for a better print result, [Online]. Available: `http://doku.radds.org/home-2/` (visited on 06/04/2018) (page 199).

[205] L. D. Woolf, "Seeing the Light : The Physics and Materials Science of the Incandescent Light Bulb", General Atomics Sciences Education Foundation, San Diego, CA, Tech. Rep., 2002 (page 202).

[206] D. Stewart. (2004). Resistivity Of Tungsten, [Online]. Available: `https://hypertextbook.com/facts/2004/DeannaStewart.shtml` (visited on 06/02/2018) (page 203).

[207] J. N. Mundy, "Electrical resistivity-temperature scale of tungsten", *Philosophical Magazine A*, vol. 46, no. 2, pp. 345–349, Aug. 1982 (page 203).

[208] Y. A. Cengel, *Heat transfer: a practical approach*, 2nd. McGraw-Hill, 2003 (pages 204, 206).

Own References

[145] G. P. Greeff and M. Schilling, "Closed loop control of slippage during filament transport in molten material extrusion", *Additive Manufacturing*, vol. 14, pp. 31–38, 2017 (pages 86, 91, 93, 98, 116, 119, 121–122).

[180] G. P. Greeff and M. Schilling, "Comparing Retraction Methods with Volumetric Exit Flow Measurement in Molten Material Extrusion", in *Special Interest Group meeting on Dimensional Accuracy and Surface Finish in Additive Manufacturing*, W. Dewulf, H. N. Hansen, R. K. Leach, J. Qian, and J. S. Taylor, Eds., Leuven, Belgium: euspen, 2017, pp. 70–74 (page 123).

[182] G. P. Greeff and M. Schilling, "Single print optimisation of fused filament fabrication parameters", *The International Journal of Advanced Manufacturing Technology*, Aug. 2018 (page 138).